FUNCTIONAL CHARACTERIZATION OF HISTONE H2A VARIANTS IN *ARABIDOPSIS THALIANA*

YELAGANDULA RAMESH
(M.Sc. Biochemistry, University of Hyderabad, India)

A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
DEPARMENT OF BIOLOGICAL SCIENCES
NATIONAL UNIVERSITY OF SINGAPORE
2013

DECLARATION

I hereby declare that the thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.
This thesis has also not been submitted for any degree in any university previously.

Yelagandula Ramesh
31 March 2014

Acknowledgments

First of all, I would like to express my sincere gratitude to my PhD supervisor, Fred, for providing me the wonderful opportunity to work in his group and for the constant support and guidance in my work without which this thesis would have not been possible. I am indebted to him for teaching me various aspects of scientific research and giving me ample freedom. His valuable help in various stages of my work, especially teaching me making preps for screening and investing his time in the daunting task of imaging is irreplaceable. I appreciate his unconditional support during ups and downs of my PhD, he is the best PI one could ever ask for.

I am thankful to our collaborators Prof. Steve Jacobsen (UCLA) and Prof. Karoline Luger (CSU) labs for their constructive suggestions, help in carrying out important experiments in completing this challenging study. I would like to specially thank Hume Stroud for his enormous amount of time and constant support he provided in analyzing genome wide data analysis.

I would like to thank my thesis committee members Toshiro, Mohan, Fumio, Ernesto and Jose Dinneny for their time and valuable suggestions. I acknowledge TLL for funding and facilities.

I am grateful to all my present lab members: Tomo, Li Jing, Thiet, Danhua, Haifeng, Nie Xin, Maru and Jeanie for creating such a wonderful lab atmosphere through their support, scientific discussions and blissful times in and out of lab, all in equal measure. I thank Nie Xin for his help in performing microarray experiments and Li Jing for his help in analyzing the data. Special thanks to Tomo for patiently listening to me and extending brotherly advice. I thank his family, Mari, Rio and Koki for wonderful experience and memorable time in the weekends. I thank all past lab members: Mathieu, Chen Zhong, Sarah, Heike, Guillaume, Peiqi and especially Pauline who taught me very patiently the molecular biology techniques, *Arabidopsis* gardening and for

her discussions and suggestions in early years of my PhD. I would like to thank other friends from TLL; Laurent, Pooja, Zimin and Emily for jovial chats and sharing experiences.

I am indebted to both Usha Vijayraghavan and L.S.Shashidhara for introducing me to science and nurturing scientific drive in me in my formative years in India. The scientific exposure I gained in their labs is invaluable; it gave me ability to swiftly adapt to new frontiers in science and immense confidence to quickly adjust to complex experiments performed in this study.

I am grateful to my friends and brothers Kanthanna, Gopanna, Ramanna and Kamini, Pooja, Usha for their care and support and for the fabulous get togethers with Indian food. I thank my friend and best senior Pushpa for happy conversations and sharing PhD experiences. I am very happy to meet Anand during my graduate years; the splendid time we spent together filled many memorable moments in my Singapore days. I thank Anil and family, Sathya and family, Amit and family, Vani and family, Kabeer and Farzana for their support and happy moments.

I am grateful to all teachers and friends whom I came across in my student life. I would like mention especially about Yadagiri sir who played an important role to take a leap in my education and my school JNVCC that gave me strong fundamental education and build courage in me to lead later life.

I am grateful to my grandmother and parents for their sacrifices in bringing me up and their bold, relentless efforts to give me the best possible education. I dedicate my thesis to them. I thank my loving brothers Nagulu and Satyam for their affection and taking care of Amma and home. I thank my very supportive and caring parents in-laws.

Last but not the least, I am grateful to my best friend and loving wife, Poornima, for her enormous support in my scientific pursuit. I appreciate her efforts in improving my writings and for our ‘at home’ scientific discussions!

Table of Contents

Summary

Histone variants play crucial roles in gene expression, genome integrity and chromosome segregation. A wide range of variants of histone H2A were identified and characterized in fungi and metazoans. While the variants H2A.X, H2A.Z and canonical H2A are conserved in all eukaryotes, species-specific H2A variants evolved in certain groups of metazoans such as the mammalian sperm cell specific H2A.B and macroH2A in vertebrates. In plants, although the existence of four classes of H2A variants was recognized from early biochemical studies, only H2A.Z was characterized functionally in *Arabidopsis*. Intriguingly, H2A.Z is confined to euchromatin in *Arabidopsis* in contrast to the profiles described in yeast and mammals. The main objective of my PhD thesis was to provide an extensive characterization of the location of all *Arabidopsis* H2A variants and to focus on the functional study of the plant specific class H2A.W.

I obtained the comprehensive genomic profiles for each class of H2A variants in *Arabidopsis*. From these maps, we find that H2A variants define different chromatin features. The chromatin at most of the gene bodies consists primarily of canonical H2A and H2A.Z, which show antagonistic distribution at the 3' and 5' end of the body respectively, that depends on the level of gene transcription. H2A.X is also distributed fairly uniform way all over the genome. Interestingly, the novel plant specific H2A variant H2A.W is exclusively localized to heterochromatin in the *Arabidopsis* genome. Consistently, in sub-nuclear localization studies, H2A.W is localized to pericentromeric heterochromatin. Functional analyses show that H2A.W is necessary and sufficient to achieve heterochromatin condensation into higher order structures. *In vitro*, H2A.W enhances chromatin condensation through a higher propensity to promote fiber-to-fiber interactions via its conserved C terminal motif. *In vivo*, elimination of H2A.W

causes decondensation of heterochromatin and conversely, ectopic expression of H2A.W promotes heterochromatin condensation. Although H2A.W recruitment is independent of heterochromatic marks H3K9me2 and DNA methylation, genetic studies show that H2A.W acts in synergy with DNA methylation associated with H3K9me2 to silence genes and transposable elements present in heterochromatin. These findings show that in plants heterochromatin regulation uses pathways distinct from metazoans and yeast, which utilize the heterochromatin protein 1 and its associated factors to silence heterochromatic elements. Moreover, I show that H2A.W conserved C-terminal like motifs are found in other variants found in metazoans, indicating that our findings will impact our understanding of heterochromatin condensation in a wide variety of eukaryotic organisms.

List Tables

List Figures

Abbreviations

Ab	Antibody
ARP6	ACTIN RELATED PROTEIN 6
ATP	Adenine Tri Phosphate
ATRX	Alpha Thalassemia Retardation Syndrome X-linked
BS-Seq	Bisulfate Sequencing
CenH3	Centromeric Histone 3
ChIP	Chromatin Immuno Precipitation
CMT2	CHROMOMETHYLASE 2
CMT3	CHROMOMETHYLASE 3
DAPI	4',6-diamidino-2-phenylindole
DBR	Doublestrand Break Repair
DDM1	DEFICIENT IN DNA METHYLATION 1
DMR2	DOMAINS REARRANGED METHYLTRANFERASE 2
DNA	Deoxyribo Nucleic Acid
FLC	FLOWERING LOCUS C
FRAP	Fluorescence Recovery After Photobleaching
gammaH2A.X	Phosphorylated H2A.X
GFP	GREEN FLUORESCENT PROTEIN
H2A.Bbd	H2A Barr body-deficient
H2A.Lap1	H2A Lack of acidic patch 1
H3K4me	Histone 3 Lysine 4 methylation

H3K9me	Histone 3 Lysine 9 methylation
H3K27me3	Histone 3 Lysine 27 methylation
HP1	HETEROCHROMATIN PROTEIN 1
INO80	INOSITOL REQUIRING 80
Kb	Kilo bases
KYP	KRYPTONITE
Mb	Megabases
MET1	METHYLTRANSFERASE 1
mH2A	macroH2A
MNase	Micrococcal Nuclease
MSCI	Meiotic Sex Chromosome Inactivation
NAD	Nicotinamide Adenine Dinucleotide
PAGE	Poly Acrylamide Gel Electrophoresis
PolyA	Poly Adenylation
rDNA	Ribosomal DNA
RFP	RED FLUORESCENT PROTEIN
RNA	Ribo Nucleic Acid
RNA-Seq	RNA Sequencing
SDC	Suppressor of DRM1 DRM2 CMT3
SDS	Sodium Dodecyl Sulfate
SIR3	SILENT INFORMATION REGULATOR 3
SUVH	SU (VAR) 3-9 HOMOLOG
SWI6	SWITCHING GENE 6

SWR1	Swi2/Snf2-related ATPase
SWR-C	SWR1 remodeling *complex*
TAIR	The Arabidopsis Information Resource
TE	Transposable Element
tRNA	Transfer RNA
TSS	Transcription Start Site
TTS	Transcription Termination Site
UV	Ultra Violet
Xist	X-inactive specific transcript

Chapter 1: General Introduction

1.1 Background

All eukaryotic genomic DNA is packaged as chromatin in the nucleus, allowing compaction of up to several meters of DNA fiber into the tiny volume of the nucleus (~15 to 20μm in diameter)(Grosberg, 2012). In addition, chromatin protects the DNA from damage and plays roles in regulation of various genomic activities. The building blocks of chromatin are five conserved basic histone proteins, H2A, H2B, H3, H4 and the linker histone H1. The basic subunit of chromatin is the nucleosome, in which 147bp of DNA is wrapped around an octamer containing a pair of each H2A, H2B, H3 and H4(Luger et al., 1997). In the constrained space of the nucleus, biological processes such as transcription, replication or DNA repair require specific mechanisms that regulate chromatin accessibility.

Post-translational modifications on the histone tails have tremendous impact on the chromatin status, for example, methylation, acetylation and phosphorylation(Jenuwein and Allis, 2001). Methylation of Histone 3 lysine 4 position (H3K4me) and acetylation of H3 and H4 tails are associated with transcriptional activation. In contrast, methylation of Histone 3 at lysine 9 (H3K9me) position is a mark for transcriptional repression. On the other hand, histone modifications also participate to chromatin mechanics. Phosphorylation of H2A tails signals for DNA double strands breaks and phosphorylation of H3 serine 10 residue is essential for chromosome condensation before mitosis(Downs et al., 2007; Prigent and Dimitrov, 2003). Histone modifications are carried out by specific “writer” complexes and are recognized by “reader” proteins(Jenuwein and Allis, 2001). Readers recruit specific complexes that regulate chromatin activity. Histone tail modifications and their causative functions are conserved in almost all eukaryotes, from

yeast to humans and higher plants(Fuchs et al., 2006). The complexity of the chromatin landscape is further enhanced by the combination of modifications present in a single nucleosome(Voigt et al., 2013).

Chromatin transcriptional activity is also regulated by the histone composition of nucleosomes(Ray−Gallet and Almouzni, 2010). In the past two decades, histone variants were discovered for core histones H2A, H2B and H3(Talbert et al., 2012). For many years, however, their function in chromatin regulation was largely overlooked. In the past few years, replacement of canonical histones with histone variants in specific regions of the genome and identification of tissue specific histone variants highlighted their important role in genome regulation(Sarma and Reinberg, 2005; Talbert and Henikoff, 2010). Understanding the dynamics of association of histone variants and their role in genomic activities is still relatively limited but marked progress was made in the past decade in functional analyses of H3 and H2A variants. I will focus my introduction on the role of H2A variants.

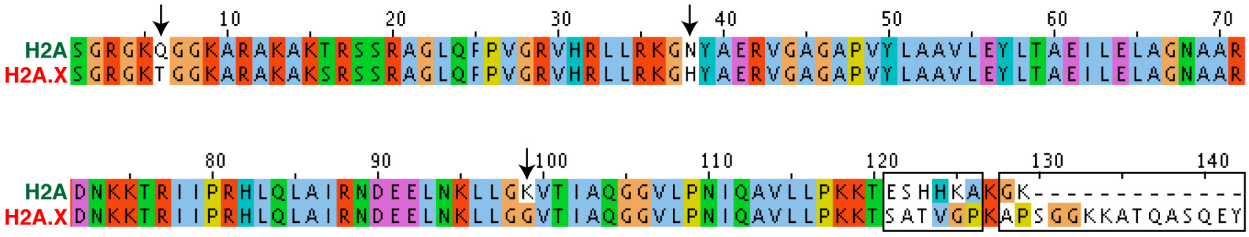

Figure 1.1. Alignment of human H2A with H2A.X
Open boxes and arrows indicate the amino acid differences between H2A and H2A.X

1.2 What are histone variants?

Histone variants are non-allelic isoforms of canonical histones with few amino acid differences (Figure 1.1). In animals, canonical histones are expressed during S-phase of the cell cycle and are incorporated into the chromatin during replication(Bönisch and Hake, 2012; Millar, 2013; Talbert and Henikoff, 2010). They are encoded by multiple copies of genes in the genome arranged in gene arrays to synchronize their expression along with cell cycle(Albig et al., 1993; Albig et al., 1997). Their genes lack introns and 3' poly adenylation signals to attain rapid high levels of expression during the S-phase to meet the demand of histone supply during replication. Canonical histones carry out the general packaging of DNA.

By contrast with canonical histones, histone variants are encoded by a single gene copy randomly placed in genome that contains introns and produces polyadenylated transcripts. Histone variants expression is independent of S-phase; they are incorporated in to the chromatin any time during the cell cycle. Except for H4 histone, variants exist for all core histones but are generally more abundant amongst the H2A and H3 classes. Attempts were made to explain why larger number of variants exists for H2A and H3 based on crystal structure of nucleosome (CHAKRAVARTHY et al., 2004). As mentioned before, in nucleosome ~147 bp of DNA is wrapped around the histone octamer, consisting a pair of each H2A, H2B, H3 and H4(Luger et al., 1997). In histone octamer, two dimers of H3-H4 are bound together by strong 4 helix bundle (HB) interactions between H3 molecules to form a H3- H4 tetramer. H3-H4 tetramer is stable in the solution even in the absence of DNA. To this tetramer, two H2A-H2B dimers bound by weak 4-HB interactions between

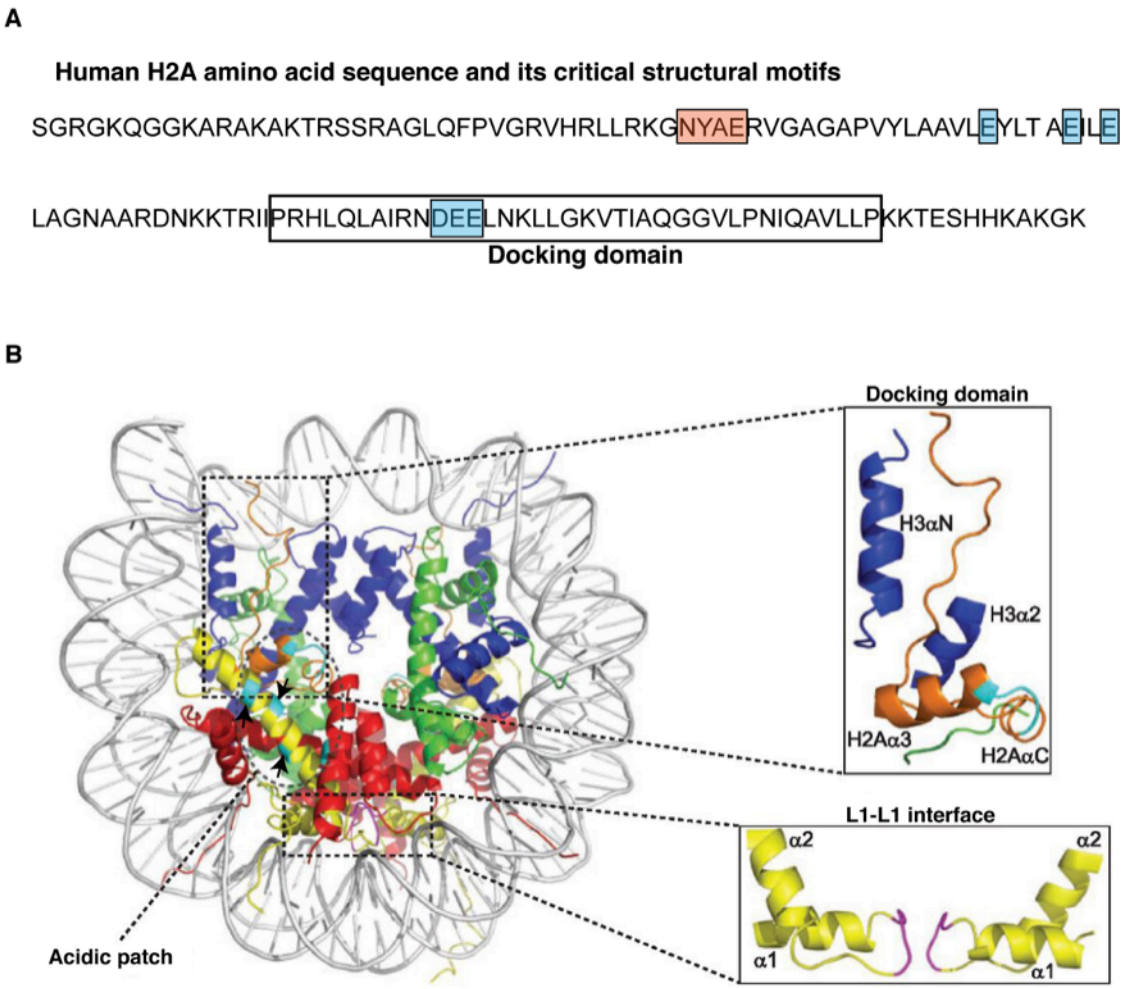

Figure 1.2. Structural motifs of histone H2A and their locations in nucleosome
A) Human H2A amino acid sequence demarcated with structural motifs. Orange box represents the L1 region, acidic patch is marked by blue boxes and open box represents the docking domain. **B)** Nucleosome crystal structure, highlighting the positions of H2A structural motifs. H2A, H2B, H3 and H4 proteins are labeled as yellow, red, blue and green respectively. DNA is labeled as grey. Arrows are pointing towards the positions of acidic patch residues in H2A. Nucleosome picture is adapted from (Bönisch and Hake, 2012).

H2B and H4. Affinities between two H2A-H2B dimers in octamer are further strengthened by interactions between L1 regions in two H2As (Figure 1.2). And also C terminal region of H2A facilitates the interaction between N terminal helix of H3 with DNA in docking region of the nucleoosme (Figure 1.2). From nucleosome structure it is clear that only H2A and H3 make homotypic interactions and, H4 and H2B never come in contact with them selves in the octamer. Intriguingly, major amino acid differences between different H3 and H2A variants are found in 4-HB region of H3 and L1 region of the H2A (Figure 1.3), which are sites for homotypic interactions. May be changes in the amino acid sequence at these interacting regions might alter the properties of nucleosomes and their functions, there by driving the divergence of H2A and H3 variants to regulate the chromatin. And also H2A C terminal tail and H3 N terminal tails make contacts with DNA entry and exit positions in the nucleosomes. H2A and H3 proximity to DNA entry-exit sites also give them opportunity to interact with histone H1 and chromatin remodelers, there by influencing the chromatin higher order organization. Consistently, in case of H2A variants, greater variation in amino acid sequence is found at their C terminal tails.

Although the definition of canonical versus variant originated form early studies, a phylogenetic view of histone families identifies that so-called canonical histones evolved from ancestral histone variants such as H3.3 and H2A.Z(Talbert et al., 2012). In the case of H3, the cell cycle controlled H3.1 class evolved independently in most phyla. Similarly, from the ancestral H2A.Z, the cell cycle controlled H2A and H2A.X are found in almost all eukaryotes. In addition the specific centromeric H3 variant (CenH3) is recruited at the centromere and is instrumental for kinetochore foundation(Earnshaw and

Tomkiel, 1992; Talbert et al., 2002). CenH3 evolves rapidly its N terminal tail and has been proposed to be instrumental in speciation in all eukaryotes(Talbert et al., 2002). In addition to these "universal" histone variants, more specific variants evolved in specific groups of organisms, reaching extreme diversity in certain parasitic groups(Talbert et al., 2012).

Histone variants affect the biochemical properties of nucleosomes, thereby influencing the property of chromatin at specific loci(Bönisch and Hake, 2012). In addition specific dynamics of nucleosome incorporation are defined for certain histone variants, which also affect reprograming of chromatin modifications(Ingouff et al., 2010). The combinatorial potential of octamers of histone variants brings further variations in the chromatin landscape. Recent advances in genome wide profiling of histone variants using chromatin immunoprecitiation followed by deep sequencing (ChIP-seq) revealed that histone variants decorate specific features of the genome. For example H2A.Z enriched at transcriptional start sites (TSS) while H3.3 enriched at 3' of gene bodies(Goldberg et al., 2010; Guillemette et al., 2005; Mito et al., 2005; Raisner et al., 2005). How this spatial ordered distribution impact chromatin functions is not well understood. From the all histone variants discovered so far, the H2A class contains the highest number of variants. These variants are identified by conserved C terminal tails and by the composition of docking domain of the histone core region (Figure 1.3). I review here H2A variants properties and functions.

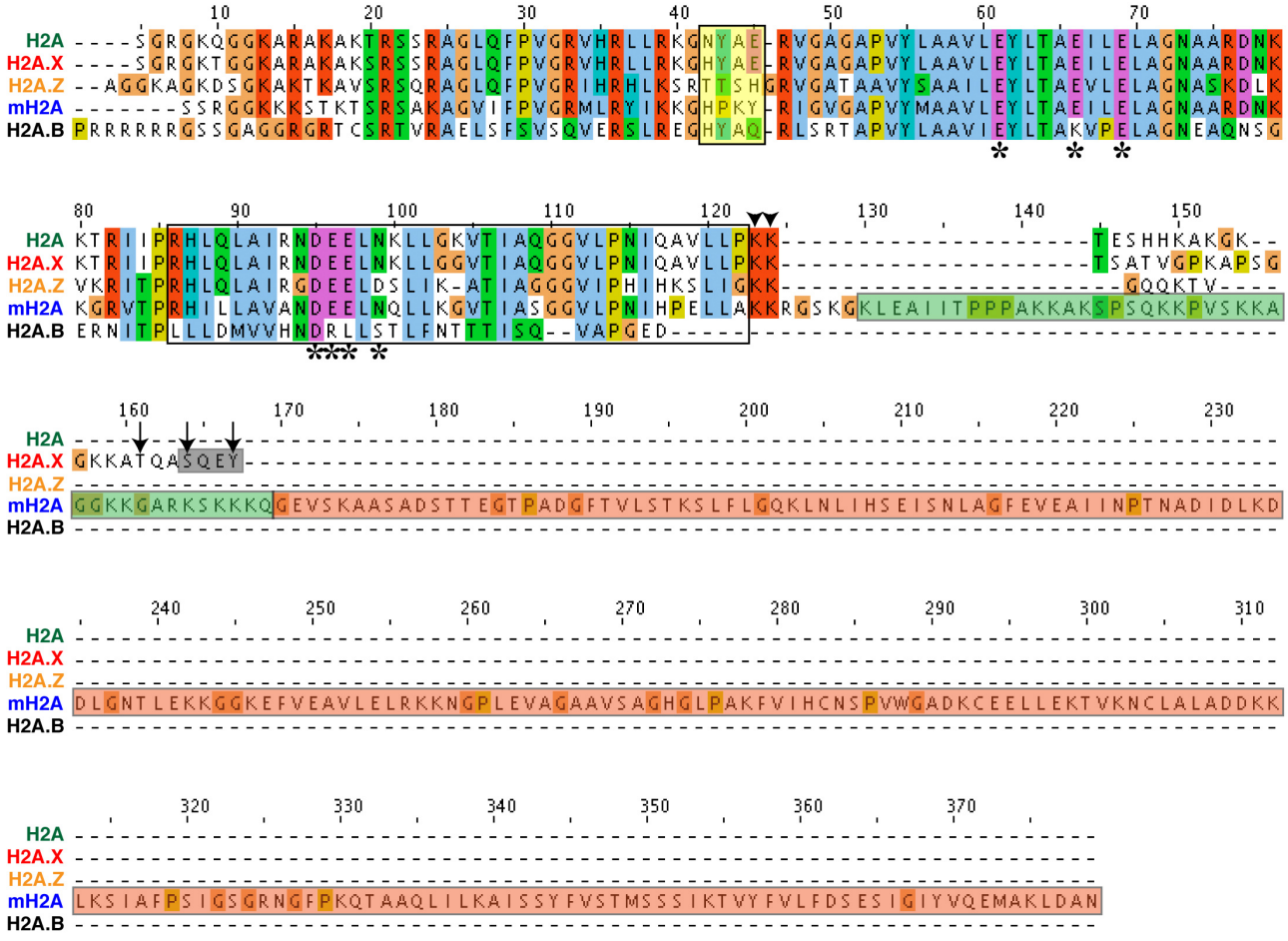

Figure 1.3. Alignment of human H2A variants and their conserved structural features
Structural features are highlighted with colored boxes. Yellow box demarcates the L1 region. Open box represents the position of docking domain. Conserved C terminal motif on H2A.X is marked with grey box. Linker region of macroH2A (mH2A) is demarcated by green box and its macro domain is highlighted with red box. Astersticks point the position of acidic patch, arrows point the phosphorylation sites on H2A.X and arrowheads point the ubiquitination sites on H2A variants. Significance of each feature is mentioned in the text describing each variant.

1.3 Histone H2A variants.

1.3.1 H2A.X

Histone H2A.X is characterized by a C terminal conserved motif Ser-Gln-(Glu/Asp)-Ø, where Ø is any hydrophobic amino acid (Figure 1.3). Apart from this conserved motif, H2A.X is mostly similar to canonical H2A (Figure 1.1). H2A.X is present in all the eukaryotes except in nematodes(Pinto and Flaus, 2010). In yeast, H2A.X conserved motif is attached to canonical H2A and there is no dedicated gene encoding for canonical H2A. Similarly, Drosophila H2Av is a H2A.Z like histone with H2A.X C terminal conserved motif(Talbert et al., 2012). Interestingly, in mammals, H2A.X transcripts produced abundantly during the S phase are devoid of PolyA tail as canonical H2A transcripts. By contrast, during the remaining part of cell cycle H2A.X transcripts are polyadenylated(Mannironi et al., 1989). This finding suggests that H2A.X loci show the properties of both replication dependent and replication independent histones genes. However, the mechanism of this differential transcript processing of H2A.X genes is not well understood.

H2A.X plays an important role in DNA damage response(Downs et al., 2007; Soria et al., 2012). Upon DNA double strand breaks, serine 139 in H2A.X is phosphorylated resulting in γH2A.X. γH2A.X can spread from several kilobases (Kb) to megabases (Mb) adjacent to the site of DNA breaks. Biochemical studies showed that γH2A.X causes loosening of nucleosomes, which is compatible with the idea that γH2A.X enables access to DNA double strand break repair machinery(Li et al., 2010). ATP driven chromatin remodelers such as SWR complex and INO80 are also important for assisting the DNA repair machinery to access DNA(Downs et al., 2007). It is proposed that γH2A.X has no role in

recruitment of chromatin remodelers but it is required for the maintenance of chromatin remodelers surrounding the DNA damage sites. Absence of H2A.X in nematodes suggests that its function is dispensable in few organisms.

γH2A.X also plays an important role in chromatin remodeling and meiotic sex chromosome inactivation (MSCI) in mouse and human(Checchi and Engebrecht, 2011; Turner, 2007). During meiosis, sister chromatids are aligned to form the synaptonemal complex. However in males, the difference between X and Y chromosome length causes partial alignment. Non-aligned regions are protected from damage by γH2A.X enrichment, which mediates transcriptional inactivation and heterochromatin formation leading to a dense structure visible in light microscopy called XY or sex body. How γH2A.X participate in this silencing is not fully understood.

In addition to serine 139 phosphorylation, phosphorylation of the tyrosine residue 142 of the SQEY motif is important in vertebrates. Interestingly, tyrosine 142 phosphorylation inhibits the phosphorylation of 139 in H2A.X, hence preventing DNA repair and promotes apoptosis by recruiting c-Jun kinase 1(Cook et al., 2009; Xiao et al., 2009). This suggests that Y142 phosphorylation plays a potential regulatory step in the activation of the DNA repair pathway.

Besides DNA repair, H2A.X plays other roles. A recent finding gave a clue about a novel role of H2A.X in chromatin remodeling in mouse embryos after fertilization(Nashun et al., 2010). In this study, authors found that high levels of H2A.X replace H2A.Z and macroH2A in mouse zygotes. Artificial incorporation of H2A.Z and macroH2A in the place of H2A.X in single cell to four cell embryos perturbs development, indicating the

importance of H2A.X enrichment. However, functional significance of H2A.X enrichment in the genome is not clear.

Our knowledge of genome wide profiles of H2A.X is still limited. In yeast, γH2A.X is enriched at heterochromatic regions(Szilard et al., 2010). Loss of γH2A.X has no effect on heterochromatin, suggesting that phosphorylation of H2A.X acts downstream of heterochromatin formation(Kitada et al., 2011). In protein coding regions, H2A.X is enriched in anti-correlation with gene expression. It also enriched on tRNA genes, rDNA, transposons and collapsed replication forks. In mammals, H2A.X and γH2A.X were mapped in two human cell types(Seo et al., 2012). In resting T cells, H2A.X shows a uniform distribution all over the genome, but upon UV irradiation γH2A.X becomes enriched over the genic regions. Where as in untreated Jurkat cells, H2A.X is enriched in subtelomeric regions and Trancription Start Sites (TSS) of genes. H2A.X in subtelomeric regions of Jurkat cells is phosphorylated. The role of H2A.X in heterochromatin maintenance and gene regulation is not well understood.

To the exception of fission yeast where γH2A.X localization to heterochromatin depends on sir3 and swi6 associated with heterochromatin maintenance(Kitada et al., 2011), mechanisms recruiting H2A.X is recruited to chromatin are unknown. A recent report states that 23 amino acids of the H2A.X C Terminal are required for its recruitment into the chromatin(Nashun et al., 2010). This suggests that recruitment signal or chaperone binding site might be present in these 23 amino acids of H2A.X. The mechanism of incorporation of H2A.X into the genome is under active search. However, the mechanism of eviction of γH2A.X from the chromatin after DBR is known, it is done by Ino80 and

Table 1

H2A variant	Expression	Localization	Function	References
H2A.X	Ubiquitous	Not well studied	DNA DSB repair, Sex chromosome inactivation	(Downs et al., 2007; Soria et al., 2012; Turner, 2007)
H2A.Z	Ubiquitous	Around TSS, In subtelomeric regions in *Saccharomyces cerevisiae*	Mostly transcription activation, Inhibits the spread of heterochromatin	(Creyghton et al., 2008; Meneghini et al., 2003; Raisner et al., 2005; Zilberman et al., 2008)
macroH2A	Ubiquitous	Inactivated X chromosome, TSS of differentiated cells	Transcription repression	(Costanzi and Pehrson, 1998a; Doyen et al., 2006; Gaspar-Maia et al., 2013; Mietton et al., 2009)
H2A.B	Testis specific	-1 nucleosome	Transcription activation	(Soboleva et al., 2012; Tolstorukov et al., 2012)

Table 1: Presenting the localization of the H2A variants in the genome and their functions.

Swr1 chaperones, which are also responsible for the incorporation of H2A.Z into the chromatin(Downs et al., 2007). This provides a good example showing the complex machinery that balances recruitment, remodeling and eviction of nucleosomes in a specific histone variant dependent manner.

1.3.2 H2A.Z

In 1980, H2A.X and H2A.Z were the first histone variants identified based on biochemical properties(West and Bonner, 1980). It is most studied H2A variant and being the ancestral H2A in eukaryotes, it is conserved from yeast to humans and plants (80% of amino acid residues)(Talbert et al., 2012). H2A.Z is only ~60% similar to canonical H2A in terms of amino acid residues. H2A.Z bears a characteristic conserved docking domain with an extended acidic patch at its C terminus (Figure 1.3). Studies in yeast, flies and humans suggest that docking domain is required for the recruitment of H2A.Z to the chromatin(Wang et al., 2011; Wratting et al., 2012). Genetic studies showed that H2A.Z is essential in flies, frogs and mouse(Bönisch and Hake, 2012). In yeast and *Arabidopsis* loss of H2A.Z also causes severe growth retardation(Carr et al., 1994; Coleman-Derr and Zilberman, 2012). These findings indicate the H2A.Z has specific functions and is non-redundant with other H2As.

In the last decade H2A.Z genome wide localization studies are extensively carried out in different species. Consistently, from all the studies it is evident that H2A.Z is enriched around the TSS(Millar, 2013). There is a species-specific difference regarding the enrichment of H2A.Z at -1 or +1 nucleosomes around the TSS. In budding yeast and humans H2A.Z is enriched at -1 nucleosomes(Creyghton et al., 2008; Raisner et al., 2005), where as in fission yeast, drosophila and in *Arabidopsis* it is enriched at +1 nucleosomes and rather depleted at -1 nucleosome(Lantermann et al., 2010; Mavrich et al., 2008; Zilberman et al., 2008). Cell type differences in enrichment of H2A.Z also observed. In mouse testis cells H2A.Z is not enriched in +1 nucleosomes as in somatic cell types(Soboleva et al., 2012).

In *Arabidopsis* H2A.Z is shown to localize in the genome in anti-correlation with DNA methylation(Zilberman et al., 2008), similar correlation was also found in human cells but not to the same extent(Valdés-Mora et al., 2012; Xiao et al., 2012). In human cells inhibition of DNA methylation by chemical treatment increased the H2A.Z occupancy at the promoters of pluripotent genes. Enrichment of H2A.Z around TSS suggests that this variant plays an important role in transcriptional regulation. H2A.Z mediated transcriptional regulation is complex. In a majority of cases it attributed to gene activation. In many eukaryotes H2A.Z is enriched at actively transcribed genes. Depletion of H2A.Z in human cells reduces the enrichment of Polymerase II at the TSS(Hardy et al., 2009). These findings indicate that H2A.Z is a transcriptional activator in the genic regions. In contrast, recent studies found that in certain loci H2A.Z acts as a transcriptional repressor. Supporting this, genome wide analysis of H2A.Z in yeast found that H2A.Z enrichment is not correlated with gene expression and accordingly, gene expression was up and down deregulated in response to H2A.Z depletion(Meneghini et al., 2003), suggesting the complexity of the role played by H2A.Z in transcriptional regulation.

Opposite effects of H2A.Z on gene transcription can result from multiple factors. Post translation modifications on H2A.Z have significant effects on gene transcription. Acetylated H2A.Z nucleosomes are enriched at active genes in comparison to inactive genes(Bruce et al., 2005). Doubly (acetylated and ubiquitinated) modified H2A.Z nucleosomes are identified on bivalent genes in human embryonic stem cells(Ku et al., 2012). Various combinations of unmodified and modified H2A.Zs in same nucleosomes participate to the regulation of transcriptional states of genes. Another level of

complexity can be brought about by the fact that H2A.Z acts in homotypic (H2A.Z/H2A.Z) or heterotypic (H2A.Z/other H2A variant) nucleosomes. *In vitro* studies showed that H2A.Z homotypic nucleosomes are more stable than canonical H2A/H2A.Z heterotypic nucleosomes and than canonical H2A homotypic nucleosomes(Bönisch and Hake, 2012; CHAKRAVARTHY et al., 2004; Park et al., 2004; Suto et al., 2000). However, *in vivo* studies in yeast showed different results. The turnover rate of H2A.Z homotypic nucleosomes is higher than that of canonical H2A nucleosomes in FRAP experiments and they are also less stable to salt treatment(Bönisch et al., 2012). These differences could be explained by the presence of factors (chromatin remodeling complexes or transcriptional machinery) that regulate nucleosome dynamics in vivo and may be preferentially attracted by the H2A.Z homotypic nucleosomes resulting in higher dynamics. The type of histone H3 variant present in the H2A.Z nuclesomes also affects its stability. *In vitro* experiments suggested that nucleosomes containing H2A.Z and H3.3 variants are least stable in different salt concentrations in comparison to H2A/H3.3 nucleosomes and canonical H2A and H2A.Z homotypic nucleosomes(Jin and Felsenfeld, 2007). Consistently, *in vivo* studies found that these hybrid nucleosomes are least stable and are enriched in nucleosome free regions(Jin et al., 2009). These results support the role of H2A.Z in transcriptional activation.

In addition to the TSS, H2A.Z is also enriched at other genomic features such as enhancer elements and heterochromatin in some species. Interestingly, recent studies found enrichment of H2A.Z at the enhancers in mouse and human cells(Hu et al., 2013), though the functional significance of this feature is not clear. H2A.Z is enriched at the subtelomeric regions in yeast but absent in the telomeric heterochromatin(Meneghini et

al., 2003). Subtelomeric enrichment of H2A.Z is important for the maintenance of chromatin boundaries i.e H2A.Z inhibits the spreading of telomeric heterochromatin in to the euchromatic regions. H2A.Z is found on pericentromeric heterochromatin in mouse and human sperm chromatin(Hammoud et al., 2009; Rangasamy et al., 2003). Intriguingly, in sperm chromatin H2A.Z is completely excluded from genic regions. In sperm, whether chromatin condensation altering the localization of H2A.Z or H2A.Z is promoting heterochromatization are few interesting aspects worth studying. H2A.Z also required for the faithful segregation chromosomes in few organisms(Rangasamy et al., 2004). In mammals H2A.Z is important for the recruitment of Heterochromatin Protein 1 (HP1) to the pericentromeric chromatin, which is important for proper organization of centromeres during the mitosis(Fan et al., 2004). These results suggest the broader role of H2A.Z in genome regulation in addition to its primary role in regulation transcription of genes.

H2A.Z deposition is carried out by ATP driven SWR1 multi protein complex (SWR-C) in yeast, fly, mammals and *Arabidopsis*(Lu et al., 2009; Mizuguchi et al., 2004). H2A.Z C-terminal docking domain is important for the SWR-C mediated recruitment(Jensen et al., 2011; Wratting et al., 2012; Wu et al., 2005). H2A.Z C terminal domain is well conserved among the different species and varies in others H2As and provides a specific scaffold for recognition by SWR-C. Depletion of SWR complex members reduces the H2A.Z levels in chromatin and also alters the localization of H2A.Z, supporting that the SWR complex mediates recruitment of H2A.Z into chromatin. SWR complex enrichment was observed around the TSS, indicating that SWR complex carries out of H2A.Z recruitment by directly binding to the chromatin(Yen et al., 2013). This binding is

mediated by a specific recognition of acetylated H3 and cis-elements in TSS(Ranjan et al., 2013; Watanabe et al., 2013). As a consequence, the genomic profile of H2A.Z is perturbed in H3 acetyltransferase mutants and insertion of promoter sequences of active genes near inactive genes causes enrichment of H2A.Z at these sites.

In addition to incorporation of H2A.Z in to the chromatin, eviction of H2A.Z is required at some gene loci to meet the biological functions of cells; the INO80 ATPase complex carries this out(Watanabe and Peterson, 2010). In the absence of INO80 complex H2A.Z becomes abnormally enriched localization over gene bodies(Yen et al., 2012). In wild type background INO80 complex enrichment at gene bodies was found in genome wide analysis(Yen et al., 2013), indicating that INO80 complex might be actively excluding the H2A.Z from gene bodies. INO80 complex enrichment was also observed at promoter sequences, suggesting a balance competition between INO80 and SWR complexes for H2A.Z dynamics.

1.3.3 macroH2A

macroH2A is a largest core histone evolved specifically in vertebrates(Talbert and Henikoff, 2010). In macroH2A, a large C terminal non histone domain is attached to a N terminal histone fold domain through a linker region (Figure 1.3) (Pehrson and Fried, 1992). Linker region in macroH2A is rich in basic amino acids similar to C terminal region of histone H1(Muthurajan et al., 2011). Macro domain of macroH2A is also exists in non-histone proteins, which are present in all kind of organisms. Macro domain has specific affinity towards nicotinamide adenine dinucleotide (NAD) and it plays wide range of functions like DNA repair, transcriptional regulation and chromatin remodeling(Han et al., 2011).

Immuno cytological studies showed that macroH2A is enriched on silenced X chromosome in mouse female cells(Costanzi and Pehrson, 1998b). This suggests that this may be involved in X chromosome inactivation. However, depletion of macroH2A has no major effects on X chromosme inactivation(Tanasijevic and Rasmussen, 2011). It may be possible that in the absence of macroH2A, other X chromosome inactivation mechanisms might compensate its function. macroH2A is recruited on to the X chromosome by recognizing the Xist RNA(Csankovszki et al., 1999). However, the chaperone that carries out the recruitment of macroH2A in to the chromatin is unknown. Interestingly, macroH2A is shown to be interact with ATRX, which is a H3.3 chaperone(Ratnakumar et al., 2012).In ATRX depleted cells macroH2A is enriched at the ectopic sites in the genome, suggesting that ATRX either evicts the macroH2A from the those ectopic sites in chromatin or inhibits the recruitment of macroH2A to these sites.

In genome wide localization analysis, macroH2A shows uniform enrichment on X chromosome consistent with the immuno cytology(Mietton et al., 2009). Outside the

inactive X chromosome, macroH2A is also enriched on euchromatic regions in larger domains where repressive mark histone 3 lisine 27 trimethylation (H3K27me3) also localized(Gamble et al., 2009), suggesting its role repression. It is also present around the TSS and on gene bodies. Macro H2A levels found to be increased in differentiated cells in comparison to pluripotent stem cells and it is shown to play role in repression of pluripotency genes there by it acts as epigenetic barrier in reprogramming (Gaspar-Maia et al., 2013; Pasque et al., 2012). Consistently, nuclear transfer experiments also shown similar results(Pasque et al., 2011). All these evidences indicate that macroH2A plays primary role in repression.

In vitro, experiments suggest that macroH2A nucleosomes are more stable than canonical H2A nucleosomes. In a recent biochemical study, macroH2A has been shown to protect the larger nucleosomal DNA from nuclease digestion in comparison to the canonical H2A nucleosome particle(Chakravarthy et al., 2012), indicating that macroH2A might be stabilizing the nucleosomes, there by repression. In another *in vitro* study, by assembling nucleosomal arrays with macroH2A its role in chromatin condensation was tested(Muthurajan et al., 2011). Where full-length macroH2A has shown similar chromatin condensation efficiency in comparison to canonical H2A. But truncated macroH2A (macrodomain deleted) surprisingly showed greater chromatin condensation efficiency. This raises the possibilities that *in vivo* through conformation changes in macrodomain or by cleavage of its macrodomain cause chromatin condensation and leading repression of its target loci. This may explain its role in the transcriptional repression of the chromatin.

Table 2

H2A variant	Effect on Nucleosome Stability	Effect on Chromatin Structure	Recruitment factor	Eviction factor	References
H2A.X	Destabilization	Not studied	Not known	INO80 complex SWR complex	(Downs et al., 2007; Soria et al., 2012)
H2A.Z	Stabilization and destabilization, Controversial evidences	Promotes chromatin folding	SWR complex	INO80 complex	(Bönisch et al., 2012; Fan et al., 2004; Mizuguchi et al., 2004; Park et al., 2004; Yen et al., 2013)
macroH2A	Stabilization	Promotes both chromatin folding and fiber-to-fiber interactions	Not known	ATRX	(Chakravarthy et al., 2012; Muthurajan et al., 2011; Ratnakumar et al., 2012)
H2A.B	Destabilization	Promotes chromatin fiber-to-fiber interactions	Not known	Not known	(Bao et al., 2004; Zhou et al., 2007)

Table 2: Presenting the effects of H2A variants on chromatin architecture and the factors, which regulate their localization in to the chromatin.

1.3.4 H2A.B

H2A.Bbd of human and H2A.lap1 of mouse were renamed to H2A.B in recent unified phylogeny based nomenclature(Talbert et al., 2012). H2A.B is a mammalian specific H2A variant, which is discovered in early years of last decade. Only in last two years, H2A.B *in vivo* expression pattern is obtained(Ishibashi et al., 2010). It is tissue specific histone with high expression in testis, accordingly, it plays essential role in spermatogenesis. H2A.B is a smallest H2A with truncated docking domain and it is only 48% similar to canonical H2A. H2A conserved acidic patch is shorter in H2A.B with replacement of few acidic amino acids with basic amino acids (Figure 1.3). Because of shorter acidic patch H2A.B shows less folding ability to make 30nm fiber *in vitro*(Zhou et al., 2007*)*.

To understand the function of H2A.B, its genome wide localization studies were carried out in human Hela cells and mouse spermitids(Soboleva et al., 2012; Tolstorukov et al., 2012). In both studies it is highly enriched around the TSS of active genes indicating the function in transcription activation. In Hela cells it is enriched also at gene bodies, exon-intron junctions, where as in mouse spermitids no such enrichment was observed. Given that significant level of endogenous transcript of H2A.B is found only in testis cells, localization pattern observed in spermitids may be realistic. Intriguingly, with depletion of H2A.B in Hela cells, results in accumulation of transcripts with retaining introns, suggesting the possible role in mRNA splicing. Since the both studies showed consistent enrichment of H2A.B at active gene promoters, it is considered to be transcriptional activator. Supporting this, in MNase treatment experiments H2A.B protects only ~120bp of DNA which is much shorter than 146bp of DNA protected by canonical H2A(Bao et al., 2004), suggesting that H2A.B makes less tight nucleosomes and easy for RNA

polymerase to unwind these nucleosomes during transcription. In another *in vitro* study, H2A.B nucleosome arrays showed more transcription activity in comparison to canonical H2A nucleosome arrays(Zhou et al., 2007). However in *in vivo*, role played by H2A.B on transcription is not clear. Interestingly, in biochemical study H2A.B also found to have higher tendency to make chromatin fiber-to-fiber interactions(Zhou et al., 2007), suggesting its role in higher order chromatin organization, though mechanism is not clear. In spermitids H2A.B might be helping in maintaining compact chromatin structure surrounding the genic regions by chromatin fiber-to-fiber interactions. It is very interesting to understand the mechanism of gene transcription in dense chromatin of spermitids.

In summery, H2A variants have large impact on genome organization, transcriptional regulation and genome integrity. H2A variants carries out wide range of functions through conserved motifs present in them, their specific localization in the genome, different combinations with other histones in nucleosomes and through posttranslational modifications present on them.

1.4 Histone H2A variants in plants

Unlike the animals in plants single copy genes randomly distributed all around the genome encode histones. Their genomic loci consist of introns and polyadenylation signal. However, canonical histones are expressed strictly during the S-phase similar to the animals(Wollmann et al., 2012). In plants, understanding the role of H2A variants in genome organization and their function is in early stages. Except for the H2A.Z, there are no significant studies carried out on other H2A variants.

1.4.1 H2A.Z role in sensing environment

Since plants are immobile, they are in a constant struggle to adapt to changing temperatures and to obtain nutrients and water. They are also vulnerable to pathogens. To overcome these problems they have developed various adaptive molecular mechanisms. Interestingly, in a recent study a novel role of H2A.Z in temperature sensing is identified in *Arabidopsis*(Kumar and Wigge, 2010). In this study, through an elegant genetic screen, authors have isolated a mutant, which behaves like it is in a high temperature adaptive state in normal conditions. By genetic mapping they found that mutant loci encode for ARP6 protein, a subunit of SWR complex, which is the chaperone for H2A.Z to recruitment it in the chromatin. By transcriptome analysis they found that most of the temperature responsive genes are upregulated in *arp6* mutant. By ChIP they showed that H2A.Z nucleosomes are enriched at the TSS of the temperature sensitive genes. With these results they came up with a model, where, when plants are in ambient temperature H2A.Z nuclesomes occupy the promoters of the temperature sensitive genes and inhibit movement of the RNA polymerase. And when plant experiences the high temperatures, H2A.Z nucleosomes will fall off and release the expression of temperature sensitive

genes. Similarly, H2A.Z also shown to be involved in the regulation of phosphate metabolism genes and pathogen response genes(March-Díaz et al., 2008; Smith et al., 2010), most likely with above-mentioned mechanism. In a very recent study indicates that H2A.Z is also enriched at the gene bodies of some genes, which show very low expression in normal conditions(Coleman-Derr and Zilberman, 2012). From transcriptome analysis they found that genes with higher H2A.Z enrichment at gene bodies have greater capacity to respond (transcriptional activation) to hostile environmental conditions and they termed these genes as "responsive genes". Role of H2A.Z in regulation of temperature sensitive genes and phosphate starvation genes may be part this responsiveness mediated by H2A.Z.

1.4.2 H2A.Z role in flowering control

In *Arabidopsis,* vegetative phase to reproductive phase or flowering transition is controlled by *FLOWERING LOCUS C (FLC)* gene(He, 2012). In vegetative phase *FLC* is constitutively expressed and represses the flowering. Once the plant reach maturity *FLC* expression is repressed by H3K27me3 and H3K9me2 posttranslational modifications on nucleosomes present at *FLC* loci. Recent study found that for proper *FLC* expression, H2A.Z presence at FLC loci is required(Deal et al., 2007). Depletion of H2A.Z leads to the early flowering phenotype. However, after flowering transition, H2A.Z is still occupied at *FLC* loci, suggesting that H2A.Z is necessary for the *FLC* loci transcription but not sufficient.

1.4.3 H2A.Z and chromatin boundaries

Genome-wide profiling of H2A.Z revealed that it is localized in anti-correlation with DNA methylation and H2A.Z is completely depleted from heterochromatin(Zilberman et al., 2008). Depletion of H2A.Z results in the hyper methylation of genes, suggesting that H2A.Z incorporation at gene loci protects them from invading from DNA methylation. However, depletion of METHYLTRANFERASE 1, which carries out CG specific methylation, has no major effect on the H2A.Z localization. What directs the H2A.Z to its targets sites in plants is not known.

Chapter 2: Materials and Methods

2.1 Plant material and growth conditions

All mutants are in Columbia-0 (Col) ecotype. *h2a.w.6* (SALK_024544.32), *h2a.w.7* (GK_149G05), and *h2a.w.12* (SAIL_667) T-DNA lines were obtained from ABRC at Ohio State University and NASC at The University of Nottingham, respectively. *drm2-2* (SALK_150863.37.35), *cmt3-11* (SALK_148381), *met1-3, ddm1*, and *kyp suvh5 suvh6* were previously characterized (Johnson et al., 2008; Jullien et al.; Vongs et al., 1993). *h2a.w.6* and *h2a.w.7* T-DNA insertions were confirmed by PCR-based genotyping. Primer sequences are described in Table 3. *Arabidopsis* plants were grown under long day conditions i.e 16 hours light and 8 hours dark.

2.2 Growth condition and crossing

Seeds were cold treated at 4°C for 4 days on soil before germination. Plants were grown in growth chamber under short day condition (8 hour day at 20°C/16 hour night at 16°C) until the rosette was formed. Flowering was induced by transferring plants to growth chamber maintained at 22°C with long day condition (16 hour-day and 8 hour-night cycles). Closed flower buds were emasculated to get clean pistil by complete removal of anthers containing immature pollen. Pistils were left for 24 hours and then manually pollinated by pollen of donor plants.

2.3 Enzymes and kits

Restriction enzymes were from New England Biolabs. Taq ploymerases used for genotype were homemade Taq polymerase from Temasek Life Sciences Laboratory and commercial DNA polymerase from GE healthcare. KOD-plus DNA polymerase from TOYOBO was used for all the cloning. Qiagen RNA miniprep kits were used for normal

total RNA extraction. Real time PCR was done with Power SYBR PCR master mix kits (ABI).

2.4 Cloning vectors

pDONR Vectors of Gateway system (Invitrogen) were used for PCR fragment cloning. Alligator2 (http://www.isv.cnrs-gif.fr/jg/alligator/vectors.html; Bensmihen et al., 2004) based gateway destination vectors and pAlli-GW43-Tnos (green color seeds selection) were developed and used for over-expression constructs. pET3a vector was a kind gift from Karoline Luger lab at Colorado state University, USA.

2.5 Bacterial and agro-bacterial strains

For standard cloning the *Escherichia coli* strains XL-blue, DH5α were used. The DB3.1 strain, which is resistant to the *ccdB* gene, was used for propagating the Gateway vectors with *ccdB* gene. *Agrobacterium tumefaciens* strain C58 was used for plant transformation. For bacterial protein expression Rosetta DE3 strain was used.

2.6 Making H2As plasmid Constructs

2.6.1 Cloning of p*H2A.X*::H2A.X-RFP for making transgenic plants

H2A.X genes a genomic fragment containing the promoter region until the codon before the termination codon was amplified with primers attB1-F-H2A.W and attB2-R-H2A.W (Table 3) using the KOD-plus-PCR kit (Toyobo). PCR fragments were cloned directionally after a BP reaction (Invitrogen) into the plasmid pDONR/Zeo (Invitrogen) to generate a pENTRY/Zeo-H2A.W entry vector. Recombination LR reactions were performed between a pAlli-2 destination vector and the pENTRY/Zeo-H2A.W vector to

generate a recombined destination plasmid with an H2A.W gene with its N-terminal fused to a gene encoding red fluorescent protein (pH2A.W::H2A.W-RFP).

2.6.2 H2As constructs for bacterial protein expression

For bacterial protein expression constructs, protein-coding sequences were amplified with primers described in Table 3 using cDNA as template. The fragments were digested with NdeI and BamHI and cloned in to pET3A vector.

All the constructs were sequenced and confirmed before used in next processes.

2.7 Plant transformation

The constructs were transformed into *A. thaliana* wild type or mutant plants by the floral dip method (Clough and Bent, 1998). In detail, respective construct was transformed in to agrobacterium(C58) by electroporation transformation method. Transformed colonies were selected on respective antibiotic selection plates. From this transforment, 250 mL was cultured overnight and collected by centrifugation. The pellet was re-suspended into 250 mL of 5% sucrose and 0.05% Silwet L-77. Siliques and opened flower were removed before plants were dipped in the solution for no more than 45 seconds. Plants were kept in moist and dark environment for 16 hours and directly resumed in long-day growth conditions. Dried seeds were harvested 1 month later. The screenings of seeds carrying GFP reporter constructs were preformed under UV by dissecting microscope. The Alligator2 origin vectors contain the GFP coding sequence driven by a seed storage protein promoter, At2S3, enabling direct visual selection.

2.8 H2A class specific antibodies

All anti-H2As class specific antibodies are affinity-purified rabbit polyclonal antibodies made by GenScript USA Inc (Piscataway, NJ) against peptides CPKKAGASKPSADE,

GGRKPPGAPKTKSVC, CKVGKNKGDIGSASQ, and KPSGSDKDKDKKKPC for H2A, H2A.W, H2A.X, and H2A.Z, respectively. Antibodies are characterized by western blot on total protein extracts for their specificity. For all samples, proteins are extracted from 10-day old seedlings except for H2A.Z.8-RFP and H2A.W.12-RFP. For H2A.Z.8-RFP protein extract is made from shoot meristem and for H2A.W.12-RFP protein extract is made from ovules. For detecting RFP and H3, anti-DsRed (Clontech, Cat.No.632496) and anti-H3 (Abcam, Cat.No. ab1791) antibodies were used.

2.9 Immunofluorescence

Immunofluorescence experiments examining the protein localization and chromatin condensation in the nuclei were performed as previously described (Jacob et al., 2009), with the following modifications. Leaves from three-week-old plants (For mutant nuclei phenotyping, leaves were collected from first generation plants. For H2A.W.7-RFP, H2A.W.12-RFP, and H3K9me2 colocalization, nuclei were isolated from roots) were fixed in 4% paraformaldehyde in TRIS buffer (10 mM TRIS pH 7.5, 10 mM EDTA, 100 mM NaCl) for 20 minutes and washed twice in TRIS buffer. Leaves were chopped in 400 microliters lysis buffer (15 mM TRIS pH 7.5, 2 mM EDTA, 0.5 mM spermine, 80 mM KCl, 20 mM NaCl, 0.1% Triton X-100) and filtered through a 40-micron cell strainer. Nuclei were concentrated by centrifugation at 500g for 3 minutes. Nuclei pellet was resuspended in 100 µl of lysis buffer and from this 10µl was air dried on cover slips. Nuclei were post-fixed in 4% paraformaldehyde in PBS for 30 minutes. Cover slips were washed three times in PBS and incubated in blocking buffer (1% BSA in PBS) for 30 minutes at 37°C. Nuclei were incubated at 4°C overnight in respective antibodies (Mouse

Anti-H3K9me2 Abcam ab1220, 1:100; Rabbit Anti-H2A.W.6, 1:100). Cover slips were washed three times with PBS and incubated with goat anti-mouse 488 and anti-rabbit 555 (Invitrogen, 1:200) for 30 minutes at 37°C. Following three PBS washes, nuclei were counterstained and mounted in Vectashield mounting media with DAPI (Vector H-1200).

2.10 Confocal microscopy

The localization pattern of H2A.W-RFP in root nuclei and immunostained nuclei were observed with laser scanning confocal microscope (Leica SPE) using 63X magnification objective. The expression pattern of H2A.W-RFP in ovules was observed using a laser scanning confocal microscopes Zeiss LSM5 Exciter and Microlamda Spinning Disk (Nikon). Images were processed using ImageJ (http://rsbweb.nih.gov/ij/). Colocalization analysis was done using auto settings in Imaris (Bitplane) software.

2.11 RNA analyses

Total RNA was prepared with RNeasy Mini kit (Qiagen, Chatsworth, CA) from ten-day-old seedlings. DNase treatment was done on 2 µg of total RNA with DNase Free Kit (Ambion, Austin, TX). RNA-seq was performed using TruSeq RNA Sample Prep Kit (Illumina) and sequenced on a HiSeq 2000 (Illumina). RNA-seq reads were processed and analyzed as previously described, where only uniquely mapping reads were retained and differentially expressed TEs and genes were defined with a four-fold cutoff and $P<0.05$ as described(Stroud et al., 2012a). Reads were mapped to the TAIR8 genome. Expression values of genes and TEs were calculated as RNA-seq reads per kilobase per million mapping reads.

For reverse transcription, 1 µg of total RNA was used to synthesize cDNA using RevertAid M-MuLV Reverse Transcriptase (Fermentas, Burlington, CA) as per

manufacturer's instructions. Semi-quantitative PCR was done using 1 µl of RT product using specific primers (Table 3). Real-time PCR assays were performed using a PCR Master Mix (Applied Biosystems, Foster City, CA) according to manufacturer's instructions using 1 µl of RT product. The specificity of the amplification product was determined by performing a dissociation curve analysis. The PCR reaction and quantitative measurements were done on 7900HT Fast Real-Time PCR System (Applied Biosystems, Foster City, CA). We performed three biological replicates, with three technical replicates for each sample. For each PCR reaction, the ΔCt was calculated using *ACT7* gene as endogenous control. Relative Quantitation values (RQ) were calculated using the $2^{-\Delta\Delta Ct}$ method ($RQ = 2^{-\Delta\Delta Ct}$).

2.12 ChIP-Seq analyses of H2As

ChIP was performed as previously described with few modifications (Wollmann et al., 2012). 2 grams of expanded leaves were ground in liquid nitrogen and fixed in 1% formaldehyde for 10 minutes. The reaction was stopped by adding 0.125 M glycine. Nuclei were extracted by filtration through Miracloth and iteratively washed and centrifuged at 2,000g. Nuclei were lysed in SDS buffer. DNA was sonicated for 10 cycles of 0.5 minute on and 1 minute off with an UCD-200TM-EX Bioruptor (Diagenode) on medium power, at 0 to 4°C. Sonicated chromatin was incubated overnight with either anti-H2As, anti-H3 (ab1791, Abcam) antibodies, or IgG (ab46540-1, Abcam). After pre-clearing, magnetic protein A-beads (Dynabeads protein A, Invitrogen) were incubated with the antibodies-chromatin mix for 3 hours. After precipitation of the beads on a magnetic rack (MagnaRack, Invitrogen) and washes with increasing stringency, DNA was eluted at 65°C and reverse cross-linked with proteinase K

(Fermentas). Immunoprecipitated DNA was treated with RNase A (Fermentas) and purified with the QIAquick purification kit (Qiagen). ChIP-seq libraries were generated as per manufacturer's instructions (Illumina). Libraries were sequenced on a HiSeq2000 (Illumina). Reads were mapped to the TAIR8 genome by allowing up to 2 mismatches, and only retaining uniquely mapping reads. Reads mapping to identical locations were collapsed in to one read and analysis performed was done as previously described(Du et al., 2012).

2.13 DNA methylation analyses

For genome-wide bisulfite sequencing (BS-Seq), plants were grown under long day condition and samples were collected from 4-week-old plants. Genomic DNA was extracted using Plant DNeasy Mini kit (Qiagen). BS-seq libraries were generated, sequenced, and analyzed exactly as previously described(Stroud et al., 2013), except that reads were mapped to the TAIR8 genome.

2.14 Alignment of protein sequences and phylogeny analysis

Histone H2A variants of *Arabidopsis thaliana* were aligned using ClustalW2 software (**www**.ebi.ac.uk/Tools/msa/**clustalw**2/*)*. Phylogenetic tree analysis of histone H2As *was* done using http://www.phylogeny.fr/version2_cgi/index.cgi (Dereeper et al., 2008). Protein sequences were obtained from http://www.uniprot.org/ Following Uniprot ID sequences were used for phylogenetic tree analysis. *Hs*_mH2A (O75367**)**, *Hs*_H2A (P04908**)**, *Hs*_H2A.X (P16104**)**, *Hs*_H2A.Z (P0C0S5), *Mm*_H2A.X (P27661**)**, *Mm*_H2A.Z (P0C0S6), *Mm*_H2A (P0C0S8), *Mm*_mH2A (Q9QZQ8), *Dm*_H2A.V (P08985), *Dm*_H2A (P84051), *Sc*_H2A (P04911), *Sc*_H2A.Z (Q12692), *Pp*_H2A.X (A9U2Y0), *Pp*_H2A (A9S4I9), *Pp*_H2A.Z (A9TYS8), *Zm*_H2A.W (P40280),

*Zm*_H2A.X (B6T543), *Zm*_H2A.Z (B4FXB2), *Zm*_H2A (B4FIA6), *At*_H2A.X (Q9S9K7), *At*_H2A (Q9LHQ5), *At*_H2A.W (Q9FJE8), *At*_H2A.Z (O23628). *Picea abies* H2A sequences were obtained from http://congenie.org/.

2.15 Protein expression and purification

Arabidopsis H2A variants H2A.1, H2A.W.6, and H2A.W.6ΔCT (C-terminal region deleted, also called CT-deletion in this paper) were expressed and purified as previously described(Muthurajan et al., 2011). All histones were prepared by the Protein Expression and Purification Facility. *Arabidopsis* H2A.1, H2A.W.6, and H2A.W.6 ΔCT were refolded together with *Xenopus* H2B to form various H2A-H2B dimers. Xenopus histone H3 and H4 are reconstituted into tetramer.

2.16 Nucleosome array experiments

2.16.1 Nucleosome array assembly

Nucleosomal arrays were reconstituted onto a DNA fragment that comprises of 12 repeats of 207 base pairs "601" positioning sequence(Lowary and Widom, 1998). A two-step reconstitution method was applied to assemble nucleosome arrays. First, different amounts of $(H3\text{-}H4)_2$ tetramers were titrated to DNA. The tetrasomes were reconstituted by step dialysis (4-6 hours each step) starting from 1 M NaCl to 0.75 M NaCl and finally 2.5 mM NaCl (in a buffer containing 10 mM Tris, pH 7.5 and 0.25 mM EDTA). The best ratio between H3-H4 and DNA for tetrasome formation was determined by sedimentation velocity AUC. Once the $S_{20,w}$ midpoint of reconstitution reached around 19S, the ratio of tetramer to DNA was fixed for the next step. Varying amounts of variant H2A-H2B dimers were added to the optimized the ratio of tetramer to DNA. The saturation levels of arrays were determined by sedimentation velocity AUC (SV-AUC). To reach a $S_{20,w}$ of

29S (characteristic of a fully saturated array), the ratio of dimer to tetrasome was around 2.0. Eco RI digestion also was be used to monitor the saturation status of nucleosome arrays as described previously(Muthurajan et al., 2011).

2.16.2 Self-association assay

Nucleosomal arrays (the equivalent of ~1 μg DNA) were combined with various concentrations of $MgCl_2$ (from 0 to 5 mM). The mixtures were incubated at room temperature for 5 minutes, and then spun at 14,000g for 10 minutes. The percentage of arrays remaining in the supernatant was determined at 259 nm, and plotted against the concentration of $MgCl_2$.

2.16.3 Folding assay by sedimentation velocity

Nucleosome arrays containing ~10 μg DNA was treated with 1 mM $MgCl_2$. SV- AUC was applied to monitor the impact of $MgCl_2$ on array folding. Sedimentation velocity experiments were done in a Beckman XL-A or XL-I ultracentrifuge at 20,000 rpm and 20 °C. The absorbance at 259 nm was scanned continuously with a redial increment of 0.003 cm. All results were processed and analyzed by UltraScanIII. By the improved van Holde-Weischet method in UltraScanIII, the diffusion-corrected integral distribution (G(s)) of sedimentation coefficients over the entire boundaries were analyzed and plotted.

Table 3

Primer sequences used in this study		
attB1-F-H2A.W.6	Cloning	GGG GAC AAG TTT GTA CAA AAA AGC AGG CT GCA GGT ATT TGA CCG TTA GTT ATC
attB2-R-H2A.W.6	Cloning	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC AGC TTT CTT TGG AGA CTT GAC TG
attB1-F-H2A.W.7	Cloning	GGG GAC AAG TTT GTA CAA AAA AGC AGG CT CGT TAA CAC CGA GCT CCA TG
attB2-R-H2A.W.7	Cloning	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC AGC CTT CTT AGG AGA TTT GGT AG
attB1-F-H2A.W.12	Cloning	GGG GAC AAG TTT GTA CAA AAA AGC AGG CT GTG TGT GAG AGT GAT GAG ATC G
attB2-R-H2A.W.12	Cloning	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC AGA TTT CTT AGG GGA TTT GGT TGC
attB1-F-H2A.X.5	Cloning	GGG GAC AAG TTT GTA CAA AAA AGC AGG CT GA CCA CAA GTG ATG GAA GAA GAG
attB2-R-H2A.X.5	Cloning	GGG GAC CAC TTT GTA CAA GAA AGC TGG GT**C** GAA CTC CTG AGA AGC AGA TCC
HTA6-TDNA-RP	Genotyping	AATACATAGTCACGCGGATCG
HTA6-TDNA-LP	Genotyping	AATCGCAAAACATCGTAATGG
HTA7-TDNA-RP	Genotyping	TTTGGAGCTTTTGAACAATGG
HTA7-TDNA-LP	Genotyping	CAAAACCCCTTTTCCAGAATC
HTA12-TDNA-RP	Genotyping	TGCAAAATCTATATGAAACGTTGC
HTA12-TDNA-LP	Genotyping	TCGAGTATCTTGCTGCTGAGG
Sail-LB	Genotyping	TAGCATCTGAATTTCATAACCAATCTCGATACAC
LBb1.3	Genotyping	ATTTTGCCGATTTCGGAAC
Gabi-LB	Genotyping	ATATTGACCATCATACTCATTGC
H2A.W.6-RT-F	Semi qRT-PCR	GAATCCACCGGAAAAGTGAAG
H2A.W.6-RT-R	Semi qRT-PCR	CTTTGGAGACTTGACTGGTG
H2A.W.7-RT-F	Semi qRT-PCR	GAGCTTTTGAACAATGGAGTCATCACA
H2A.W.7-RT-R	Semi qRT-PCR	GCAGAAGCTTTCTCCGCTTGAG
H2A.W.12-RT-F	Semi qRT-PCR	GCTGGAAGAAGAAGTGGTG
H2A.W.12-RT-R	Semi qRT-PCR	CCTTTGATGGTGATTTGGGT
ACTIN 7-qPCR-R	RT-qPCR	TCGTGGTGGTGAGTTTGTTAC
ACTIN 7-qPCR-F	RT-qPCR	CAGCATCATCACAAGCATCC
SDC-qPCR-R	RT-qPCR	AATGTAAGTTGTAAACCATTTGAACGTGACC
SDC-qPCR-F	RT-qPCR	CAGGCATCCGTAGAACTCATGAGC
soloLTR-qPCR-R	RT-qPCR	AACTAACGTCATTACATACACATCTTG
soloLTR qPCR-F	RT-qPCR	AATTAGGATCTTGTTTGCCAGCTA
H2A.W.6-Fwd-NdeI	Cloning	GGGAATTCCAT ATG GAA TCC ACC GGA AAA GTG AAG AAA G
H2A.W.6-Rev-BamHI	Cloning	CGC GGA TCC GCG TTA AGC TTT CTT TGG AGA CTT GAC TGG
H2A.W.6-CT-del-Rev-BamHI	Cloning	CGC GGA TCC GCG TTA CTT CTT AGG CAA TAG AAC AGA G
H2A.1-Fwd-NdeI	Cloning	GGGAATTCCAT ATG GCT GGT CGT GGA AAA ACT CTT GGT TCC GGT G
H2A.1-Rev-BamHI	Cloning	CGC GGA TCC GCG CTA ATC TTC CTG AGG CTT TGA AG

Chapter 3: Genome-wide Profiling of Histone H2A variants in *Arabidopsis thaliana*

3.1 Background

The aim of this study is to investigate the histone H2A variants genomic occupancy and their functions in *Arabidopsis thaliana*. As discussed before in the introduction, a wide range of histone variants are found and characterized in yeast, flies and metazoans. Universal variants like H2A.X and H2A.Z are conserved in all eukaryotes. Also species-specific variants such as H2A.B and macroH2A are evolved in mammals and vertebrates respectively. In plants, when I initiated my PhD, the study of histone variants was at its early stages. Only H2A.Z was characterized in *Arabidopsis*, and intriguingly the first profiles showed that it is enriched at TSS of genes in euchromatin (Zilberman et al., 2008). H2A.Z plays a role in flowering, temperature sensing and stress response (Coleman-Derr and Zilberman, 2012; Deal et al., 2007; Kumar and Wigge, 2010). It was not known which H2A(s) participate(s) in building the chromatin over gene bodies, heterochromatin regions and other genomic features. The function of other H2A variants was also not well understood. Here, I present a study that aims to answer these questions. I generated the genome wide localization map for each type of H2A variant in the *Arabidopsis* genome and carried out the detailed functional characterization of a novel H2A.W class.

3.2 H2A variants in *Arabidopsis* are classified into 4 classes based their C terminal conserved regions

The *Arabidopsis* genome comprises thirteen genes encoding histone H2A variants, which are distributed randomly in the genome. All the genes have a single intron and their transcripts are polyadenylated. Aligning the H2A proteins suggest that histone core regions in between them are highly conserved except in L1 region and docking domain

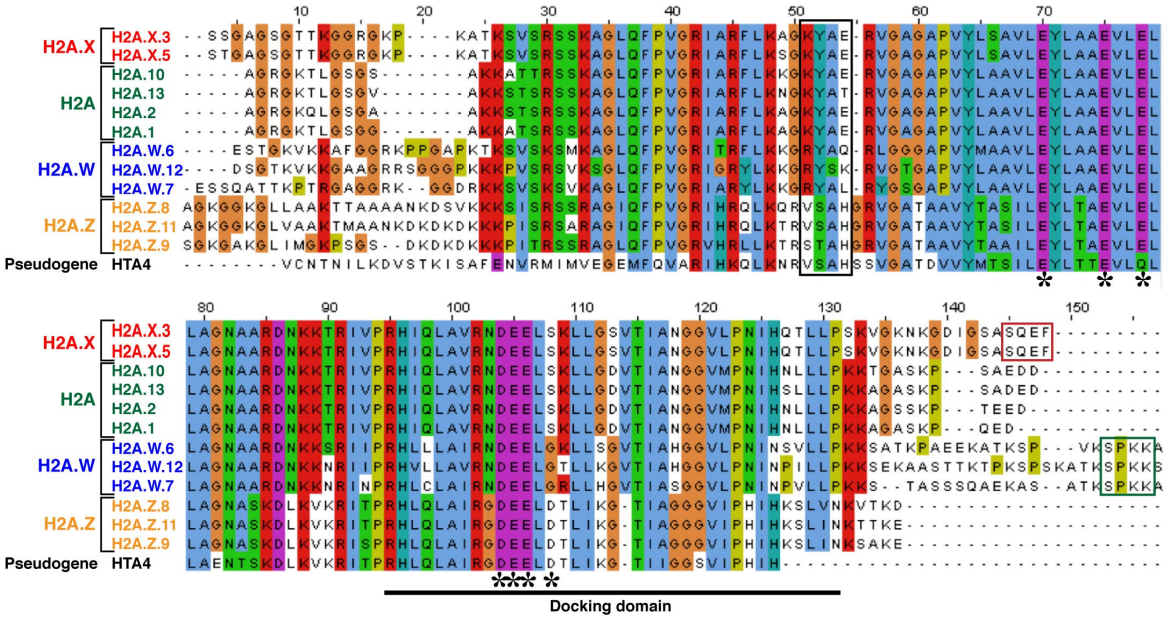

Figure 3.1. Alignment of histone H2A variants of *Arabidopsis thaliana*
H2A variants alignment is generated using ClustalW2 program. H2A variants are classified into four classes based on their conserved C-terminal tails and docking domain. Open boxes represent conserved motifs and asterisks represent acidic patch.

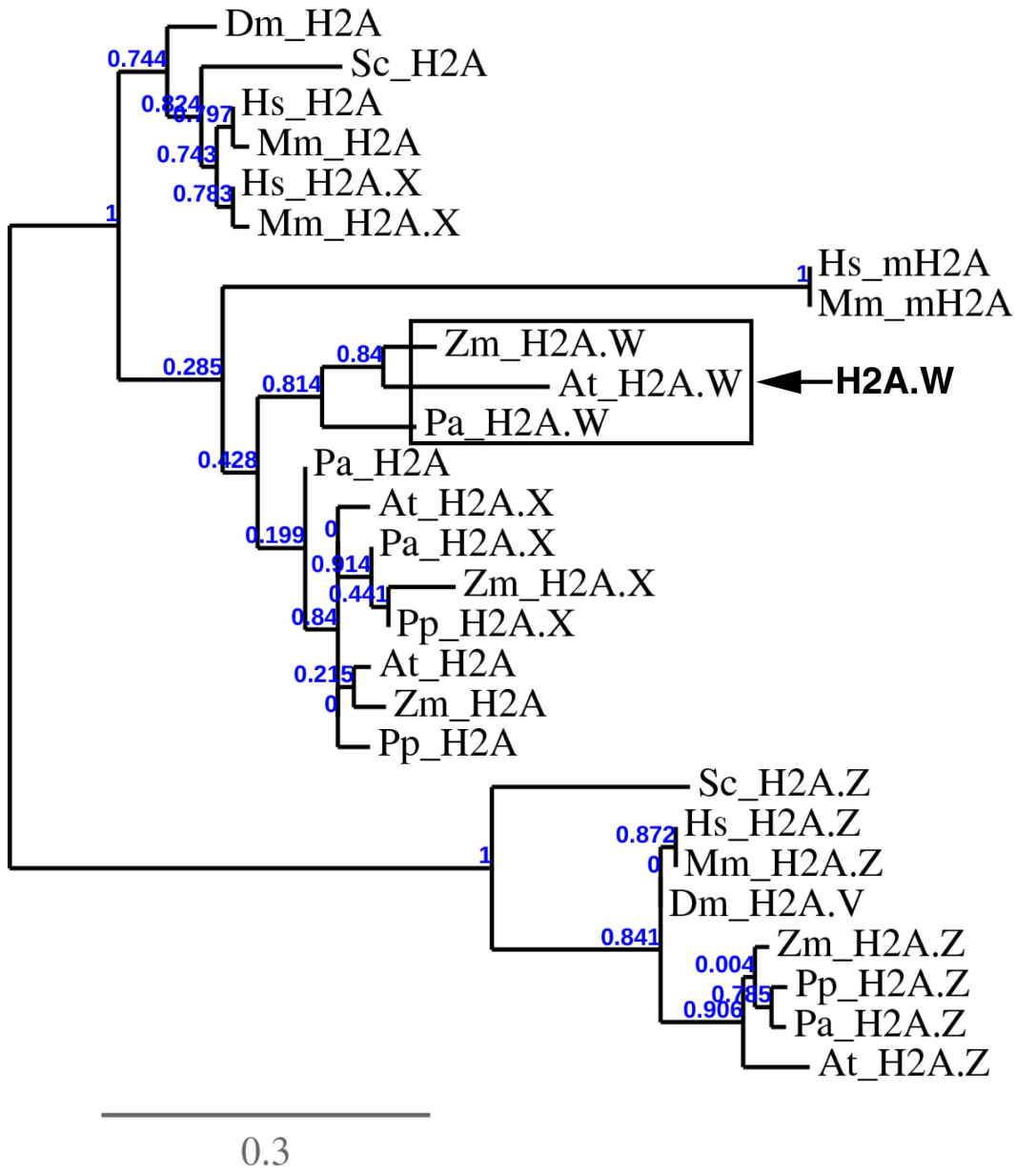

Figure 3.2. Phylogenetic tree of H2A variants
Where At: *Arabidopsis thaliana*, Dm: *Drosophila melanogaster*, Hs: *Homo sapiens*, Mm: *Mus musculus*, Pa: *Picea abies,* Pp: *Physcomitrella patens*, Sc: *Saccharomyces cerevisiae* and Zm: *Zea Mays.* Branch supporting values are colored in blue.

(Figure 3.1). Large variation is found in their N terminal and C terminal tails. They are grouped into four classes according to their C-terminal conserved motifs (Figure 3.1) (Talbert et al., 2012). In Arabidopsis, the canonical H2A class, referred to as H2A class thereafter, consist of four histones with acidic residues in the C-terminal tail. Two histones form the H2A.X class identified by the conserved motif SQEF (Figure 3.1) both in plants and animals (Figure 1.3). The H2A.Z class comprises three members. H2A.Z share the shortest length in aa residues compared with other H2A variants and are characterized in their core domain by an extended acidic patch and a conserved docking domain. The Arabidopsis gene encoding the H2A.Z-like HTA4 is a pseudogene (Coleman-Derr and Zilberman, 2012). Interestingly, the Arabidopsis genome comprises three genes encoding H2.A variants characterized by the C terminal motif SPKK. These proteins represent the novel class H2A.W. This class is conserved and found only in all seed bearing plants i.e Angiosperms and Gymnosperms (Figure 3.2).

3.3 Generation of H2As class specific antibodies

In order to obtain genome-wide profiles of each H2A variant class, we performed Chromatin Immuno Precipitation (ChIP) of at least one representative member of each class. First, I obtained specific antibodies for each class of H2A variant. To generate antibodies I used peptide sequences in N terminal or C terminal tails outside of the conserved histone H2A core region. For each H2A class, I selected a specific peptide sequence, which is least similar to other members of the same class. Using these peptides I raised the antibodies in rabbits. All the antibodies were peptide affinity purified. Antibodies specificity is tested in western blots (Figure 3.3).

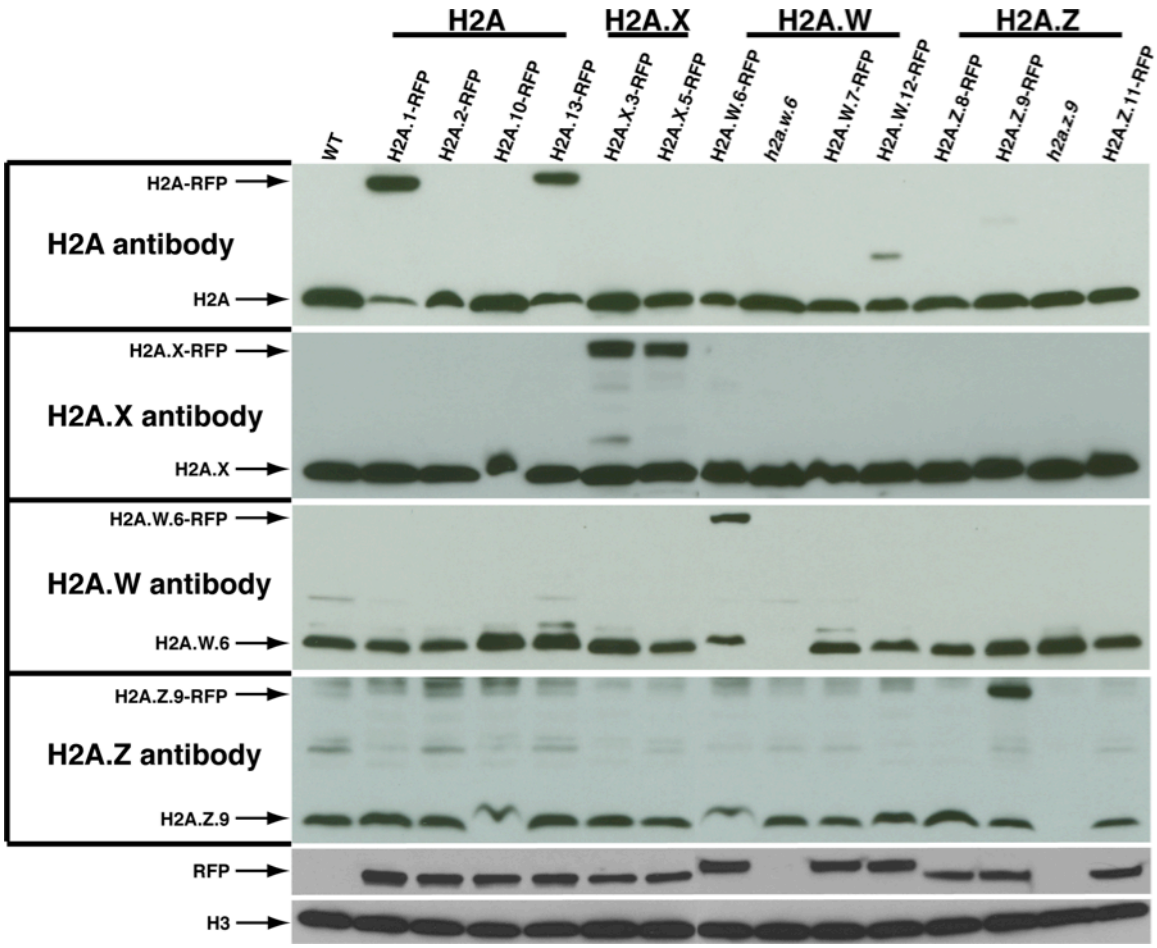

Figure 3.3. Western blot confirming the specific binding of each antibody. The antibodies are tested on whole cell protein extracts. In H2A-RFP transgenic protein samples both respective wild type H2A and H2A-RFP fusion proteins are present. The abundantly transcribed H2A.1 and H2A.13 are recognized by the anti-H2A antibody. The anti-H2A.X antibody recognized both members of the family. The anti-H2A.Z antibody is specific for H2A.Z.9. The anti-H2A.W is specific for H2A.W.6 that is most highly expressed amongst all *H2A.W* genes. No cross reactivity is observed to RFP fusion proteins of other class members, confirming specificity of the antibodies. For H2A.Z and H2A.W no band was observed in respective T-DNA knocks out lines, further confirming the specificity of anti-H2A.Z and anti-H2A.W antibodies.

3.4 Genome-wide profiling of H2A variants

In order to obtain genomic profiles of H2A variants I performed chromatin immunoprecipitation for each class of H2A variant using their specific antibodies and purified the bound DNA. Using this bound DNA, I constructed DNA sequencing libraries using illumina reagents. These libraries were subjected to deep sequencing (ChIP-seq) in UCLA sequencing facility in collaboration with Steve Jacobsen Laboratory. The sequences obtained were aligned to TAIR 8 genome leading to a comprehensive localization map for all H2A variants (Figures 3.4, 3.5 and 3.6). H2A.X covered the entire genome uniformly (Figures 3.4 and 3.5), in keeping with its role in DNA repair (Downs et al., 2007; Soria et al., 2012). H2A showed a relative depletion at pericentromeric chromatin as well as over transposable elements (TEs) and islands of H3K9me2 present in chromosome arms outside pericentromeric heterochromatin (Figures 3.4, 3.5B and 3.5C). Both H2A and, to a lesser extent H2A.X, were enriched over gene bodies (Figure 3.5A).

In contrast with the rather uniform profiles of H2A and H2A.X over gene bodies, H2A.Z was enriched specifically in the 5' end of genes (Figure 3.5A) and was highly depleted over all heterochromatic features (Figures 3.5B and 3.5C), as reported previously. H2A.Z enrichment around TSS is similar to that observed in humans, *Drosophila* and yeast (Hu et al., 2013; Leach et al., 2000; Raisner et al., 2005). Unlike all other H2A variants, H2A.W was strongly depleted from gene bodies (Figure 3.5A). H2A.W was specifically enriched in pericentromeric heterochromatin, TEs, and islands of H3K9me2 (Figures 3.4 and 3.5B-D). The Pearson correlation between H3K9me2 and H2A.W within TEs across

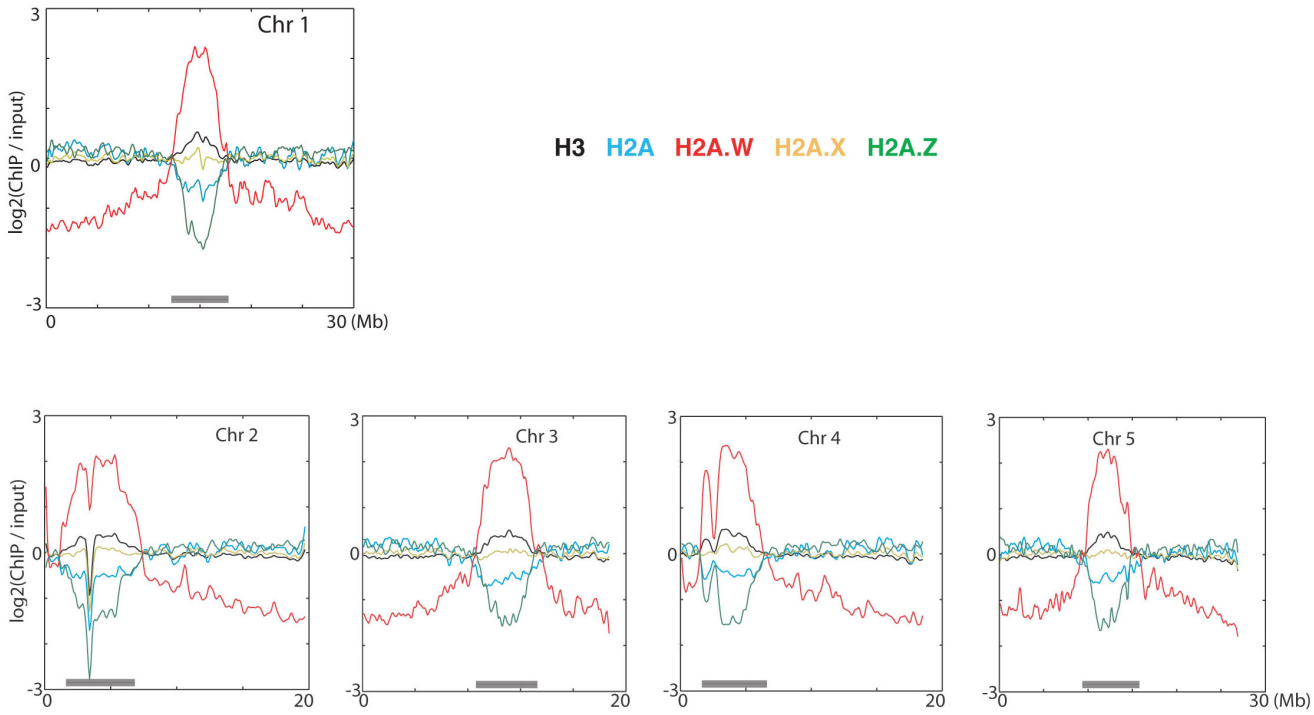

Figure 3.4. Genome-wide profiles of H2A variants in *Arabidopsis*
(A) Distribution of ChIP-seq reads along five Arabidopsis chromosomes for each H2A variant and H3 relative to input DNA. Grey bar at the bottom indicate pericentromeric heterochromatin.

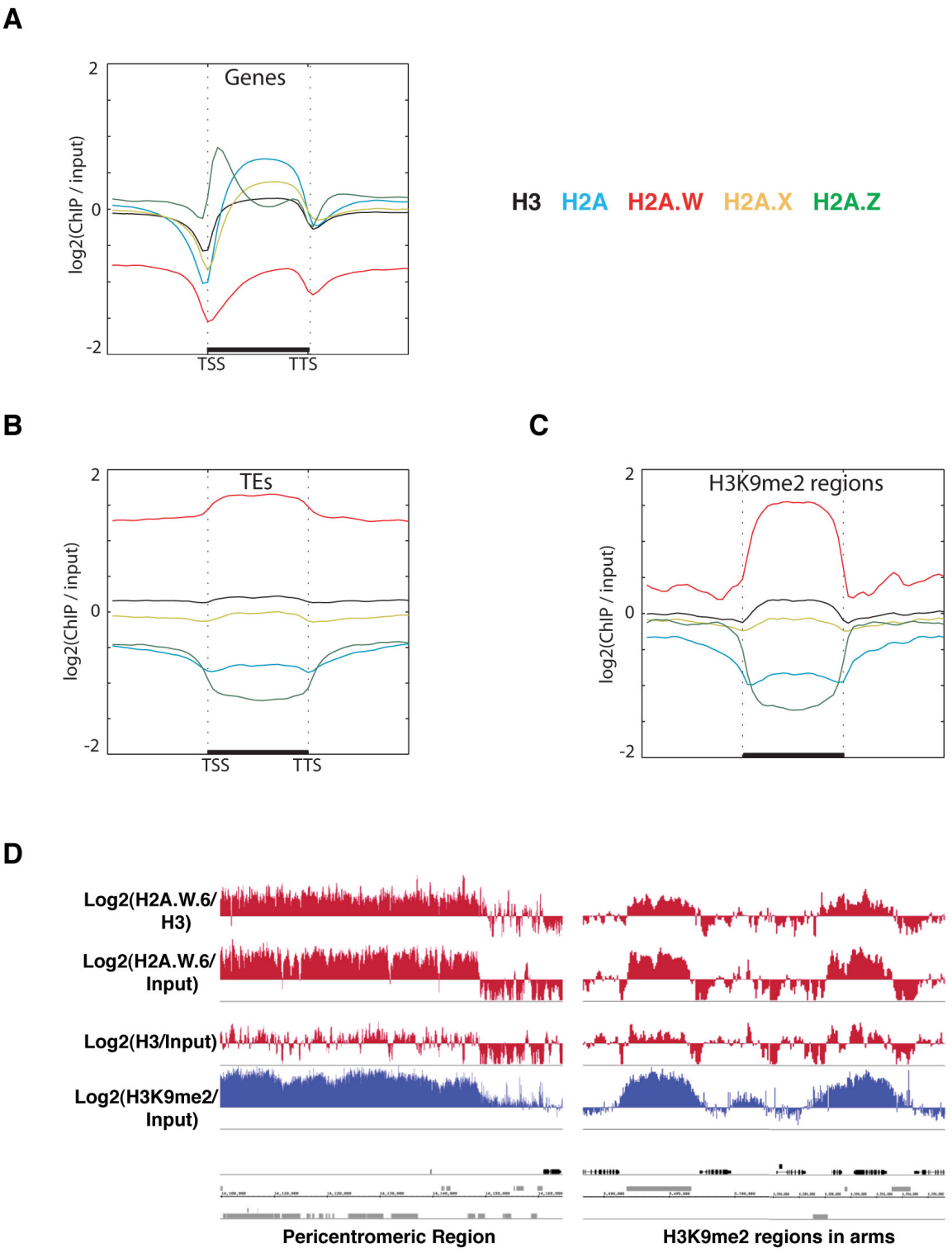

Figure 3.5. H2A variants localization pattern in different genomic features
(A) Average enrichment of H2As and H3 over genes delimited by transcription start site (TSS) and transcription termination site (TTS). (B) Average enrichment of H2As and H3 over transposable elements (TEs). (C) Average enrichment of H2As and H3 over defined

H3K9me2 regions in chromosomal arms. In (A to C) the middle region marked by the horizontal bar represents the body of the elements. The left and right flanking regions represent the 5' upstream and the 3' downstream regions, respectively, and were scaled such that their lengths are same as the middle region. D) Genome browser views of H2A.W.6 enrichment over pericentromeric region and H3K9me2 region in chromosomal arms. Black bars represent genes grey bars represent TEs.

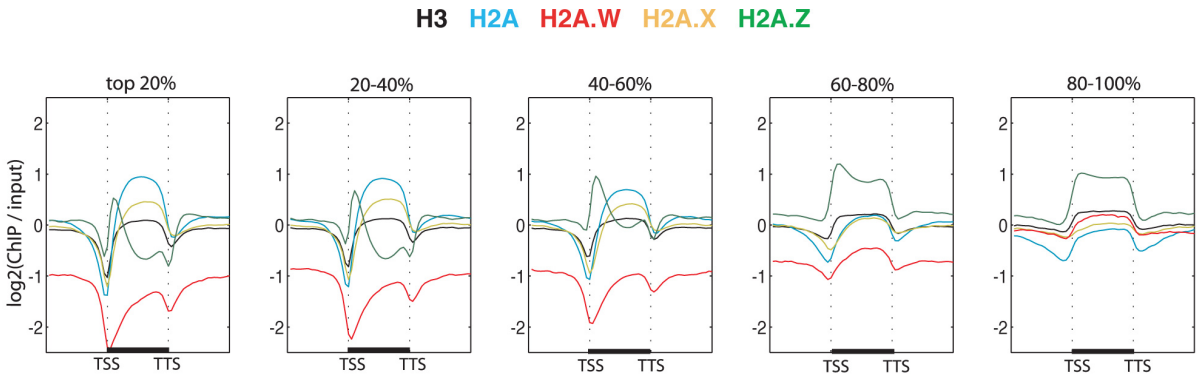

Figure 3.6. Localization of H2A variants on gene bodies in relation with gene expression levels

Distribution of ChIP-seq reads along the gene loci classified according to their expression levels for each H2A variant and H3 relative to input DNA.

the genome was 0.7, indicating a very high degree of overlap between H2A.W and H3K9me2. In addition we analyzed whether enrichment of each H2A variants depends on the level of gene expression. Comparing genes expressed from high to low levels we observed a gradual depletion of H2A and H2A.X and in turn an enrichment of H2A.Z over gene bodies (Figure 3.6). Overall I conclude that antagonistic localizations of H2A.Z and H2A.W distinguish active and inactive domains of the chromatin, respectively. My results further suggest that H2A.W is associated with transcriptional silencing.

Chapter 4: Functional Characterization of a Novel Heterochromatin Specific H2A Variant H2A.W

4.1 Background

Chromatin is organized into transcriptionally active euchromatin and more compact inactive heterochromatin (Heitz, 1928). In somatic cells of *Arabidopsis*, heterochromatin can be visualized as dense structures called chromocentres in DAPI staining (Fransz et al., 2006). In *Schizosaccharomyces pombe* and mammals heterochromatin is demarcated by histone H3 lysine 9 methylation (H3K9me3) (Jenuwein and Allis, 2001). Heterochromatin protein 1(HP1) binds to the H3K9me2 and recruits other protein complexes and aids in heterochromatin condensation and transcriptional silencing (Grewal and Elgin, 2007). In *Arabidopsis,* heterochromatin is similarly marked by H3K9me2, which is deposited by the histone methyltransferases KRYPTONITE (KYP), SUVH5 and 6 (Bernatavichute et al., 2008; Jackson et al., 2002; Malagnac et al., 2002). In addition, the *Arabidopsis* CHROMOMETHYLASE 3 (CMT3) binds H3K9me2 (Du et al., 2012; Lindroth, 2001) and methylates cytosine residues at CHG sites (where H is A or C or T). In turn CHG methylation recruits KYP in an action that completes a positive feed-back loop.

In addition to CHG methylation, heterochromatin is also marked by DNA methylation in CG and CHG contexts. CG DNA methylation is carried out by METHYLTRANSFERASE 1 (MET1) (Law and Jacobsen, 2010). DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2) and the recently discovered CHROMOMETHYLASE 2 (CMT2) perform methylation in CHH contexts (Haag and Pikaard, 2011; Zemach et al., 2013). DNA methylation and H3K9me2 silence genes and transposable elements (TEs) in *Arabidopsis* heterochromatin (Law and Jacobsen, 2010).

Although H2A.Z and H3.3 are predominantly linked to transcriptional activation, they are also associated with heterochromatin maintenance. In yeast, H2A.Z is also enriched in subtelomeric regions and prevents spreading of telomeric heterochromatin into euchromatic regions (Meneghini et al., 2003). Similarly, in humans H3.3 is localized to telomeric heterochromatic regions (Goldberg et al., 2010; Wong et al., 2009). In contrast, in *Arabidopsis* H2A.Z and H3.3 are excluded from heterochromatin. H2A.Z is enriched at TSS of expressed genes and also enriched on gene bodies of responsive genes (Coleman-Derr and Zilberman, 2012; Zilberman et al., 2008). Similarly, H3.3 enrichment is confined to gene bodies and is correlated positively with gene expression (Stroud et al., 2012b; Wollmann et al., 2012).

In contrast to animal and yeast heterochromatin marked by HP1, in Arabidopsis the HP1 homolog LHP1 is completely absent in the heterochromatin. Similarly, the histone variants H2A.Z and H3.3 appear to define exclusively located to euchromatin. These observations indicate that in plants pathways regulating plant heterochromatin are distinct from that defined in yeast and animals. Novel mechanisms hat define and regulate heterochromatin likely exist as suggested by the specific heterochromatic localization of H2A.W in my genome-wide profiling study. I hypothesize that H2A.W might be regulating the heterochromatin through a novel mechanism.

4.2 H2A.W indexes heterochromatin

H2A.W carries an extended C terminal tail rich in positive amino acids with a conserved SPKK (in general T/SPXK) motif (Bönisch and Hake, 2012). As mentioned before, the H2A.W variant class evolved and is highly conserved in seed bearing plants (Figure 3.2).

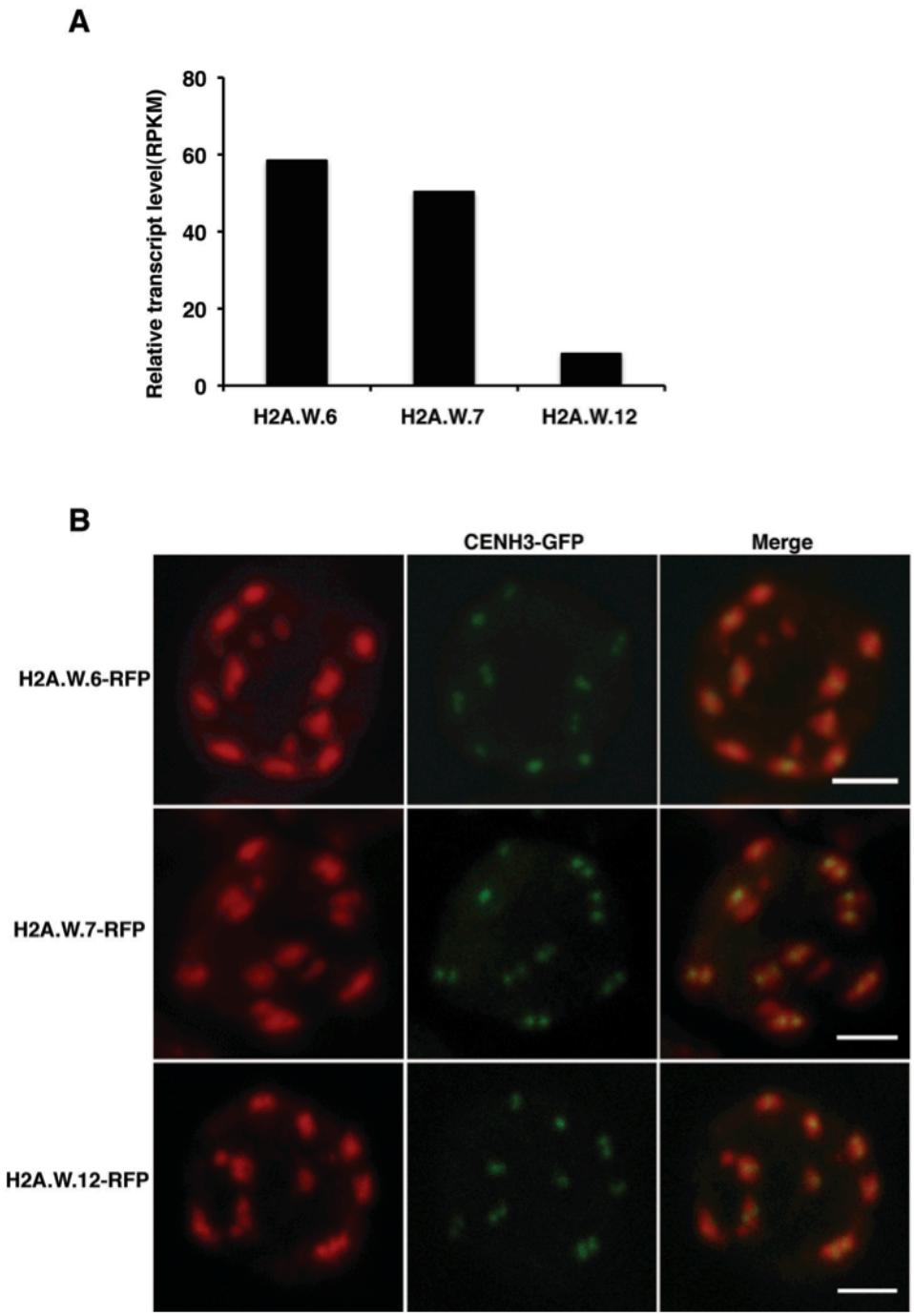

Figure 4.1. Expression levels of H2A.W transcripts and their proteins localization pattern in somatic cell nuclei of *Arabidopsis*

(A) Expression levels of *H2A.W.6*, *H2A.W.7*, and *H2A.W.12* transcripts. Reads per kilobase per million reads (RPKM) values were obtained from RNA-seq data of wild type. (B) Confocal sections of a root nuclei of transgenic plants expressing p*CENH3::CENH3-GFP* and *pH2A.W.X::H2A.W.X-RFP*. In confocal sections, scale bars represent 2 μm.

In *Arabidopsis* the H2A.W family comprises three proteins H2A.W.6, H2A.W.7, and H2A.W.12, which are encoded by the genes *HTA6*, *HTA7*, and *HTA12*, respectively (Figure 3.1). *HTA6* is expressed at higher levels than the other two H2A.W encoding genes (Figure 4.1A) and I used an antibody that recognizes H2A.W.6 specifically to perform ChIP-seq analyses. In order to investigate whether the properties of H2A.W.6 revealed by ChIP-seq analysis were representative of the entire H2A.W class, I analyzed the localization of each member of the H2A.W family *in vivo*. I fused the RED FLUORESCENT PROTEIN (RFP) to each gene encoding H2A.W and expressed the fusion protein under the control of its respective promoter in transgenic plants that were also expressing the CENTROMERIC HISTONE 3 (CenH3) fused to the GREEN FLUORESCENT PROTEIN (GFP). CenH3 is located at the centromere of each chromosome and is surrounded by pericentromeric chromatin that forms heterochromatin domains called chromocenters (Fang and Spector, 2005; Talbert et al., 2002). Each H2A.W-RFP localized primarily to domains corresponding to the ten chromocenters, each containing a CENH3-GFP patch in its center (Figures 4.1B). In ChIP-seq analysis, H2A.W.6 co-localizes with the heterochromatin mark H3K9me2. Consistently, immunostainings also show perfect co-localization of H2A.W.6 with H3K9me2 (Figure 4.2A). Similarly, H2A.W.7-RFP and H2A.W.12-RFP also co-localize with H3K9me2 (Figures 4.2B-C). In conclusion, all members of the H2A.W family are tightly associated with heterochromatin in somatic cells.

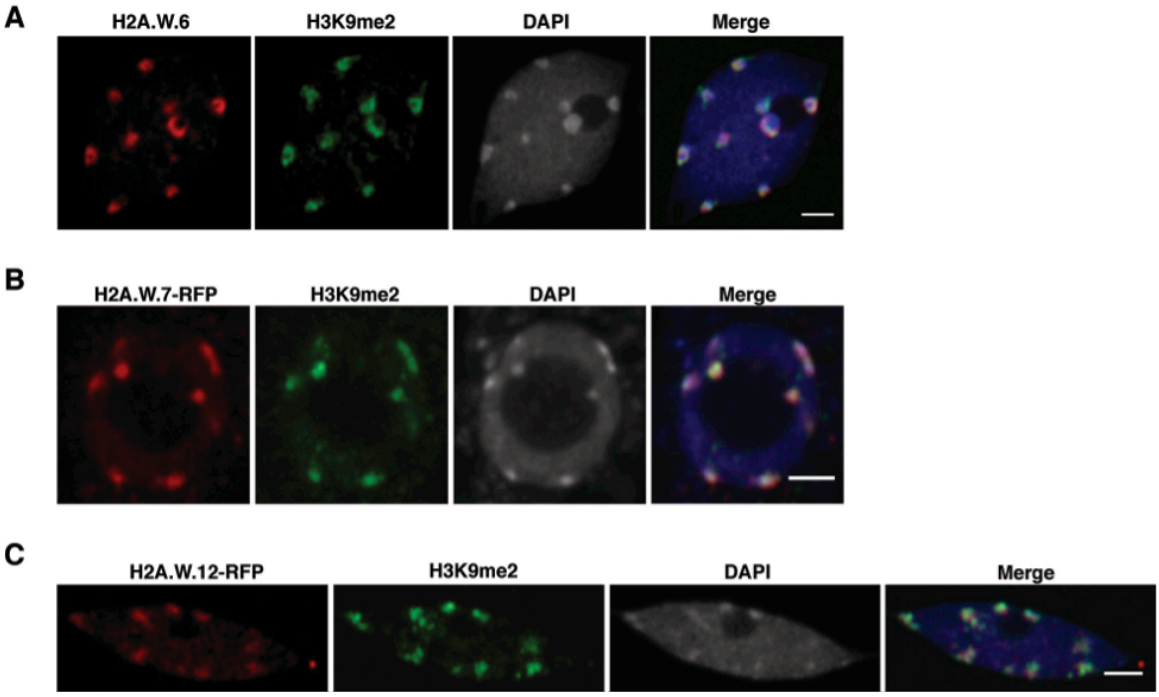

Figure 4.2. Immuno localization of H2A.W proteins in the somatic cell nuclei of *Arabidopsis*
(A) Confocal section of a leaf nucleus immunostained for H3K9me2 and H2A.W.6 with DAPI counterstaining (B) Confocal section of a leaf nucleus immunostained for RFP and H3K9me2 with DAPI counterstaining. (C) Confocal section of an ovule integument nucleus immunostained for RFP and H3K9me2 with DAPI counterstaining. Scale bars represent 2 μm.

4.3 Recruitment of H2A.W is independent of H3K9me2 and DNA methylation

The co-localization between H2A.W and H3K9me2 suggested a link between the mechanisms responsible for deposition of these two heterochromatic features. I asked whether the H2A.W recruitment is dependent on H3K9me2. Preventing H3K9me2 deposition by mutating the three methyltransferases KYP, SUVH5, and SUVH6 did not significantly alter the pattern of H2A.W.6 (Figure 4.3). Similarly, H2A.W.6 localization was not perturbed significantly by the absence of CMT3, that methylates cytosine residues in CHG contexts at heterochromatic loci marked by H3K9me2 (Figure 4.3). Since H2A.W localization does not depend on H3K9me2 and the associated CHG methylation, I tested the influence of other heterochromatic pathways responsible for DNA methylation. The loss of the *de novo* DNA methyltransferase DRM2 did not affect H2A.W nor heterochromatin organization (Figure 4.4A). In contrast, heterochromatin was partially dispersed in the absence of the CG methyltransferase MET1 and the chromatin remodeling protein DECREASED DNA METHYLATION 1 (DDM1) that contribute to maintenance of DNA methylation and silencing of heterochromatic loci (Figure 4.4A) (Vongs et al., 1993; Zemach et al., 2013)(Jeddeloh et al., 1999). In spite of the change of heterochromatin condensation in *met1* and *ddm1* genetic backgrounds, H2A.W.6 still co-localized with dispersed heterochromatin marked with H3K9me2 (Figures 4.4A-B). These results suggest that the recruitment or maintenance of H2A.W to heterochromatin do not depend on DNA methylation or H3K9me2.

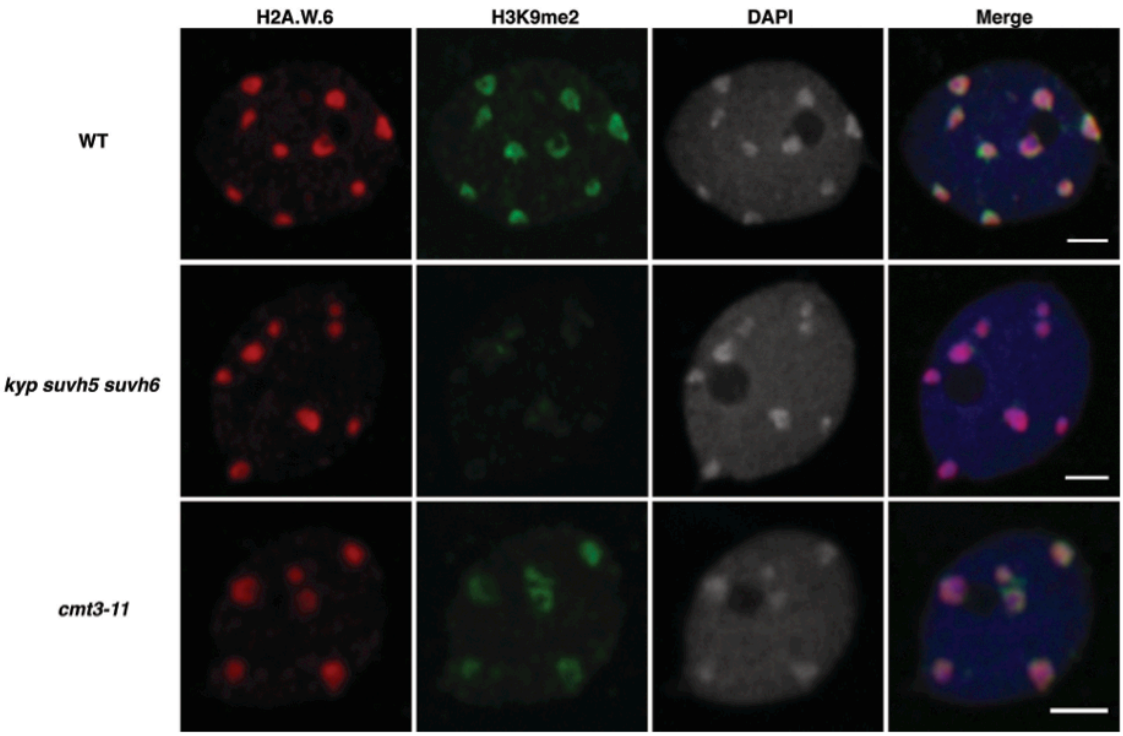

Figure 4.3. Recruitment of H2A.W.6 to chromatin is independent of H3K9me2 and CMT3 mediated DNA methylation

Confocal sections of leaf nuclei immunostained for H3K9me2 and H2A.W.6 with DAPI counterstaining in genetic backgrounds WT, *kyp suvh5 suv6*, and *cmt3-11*. Scale bars represent 2 μm.

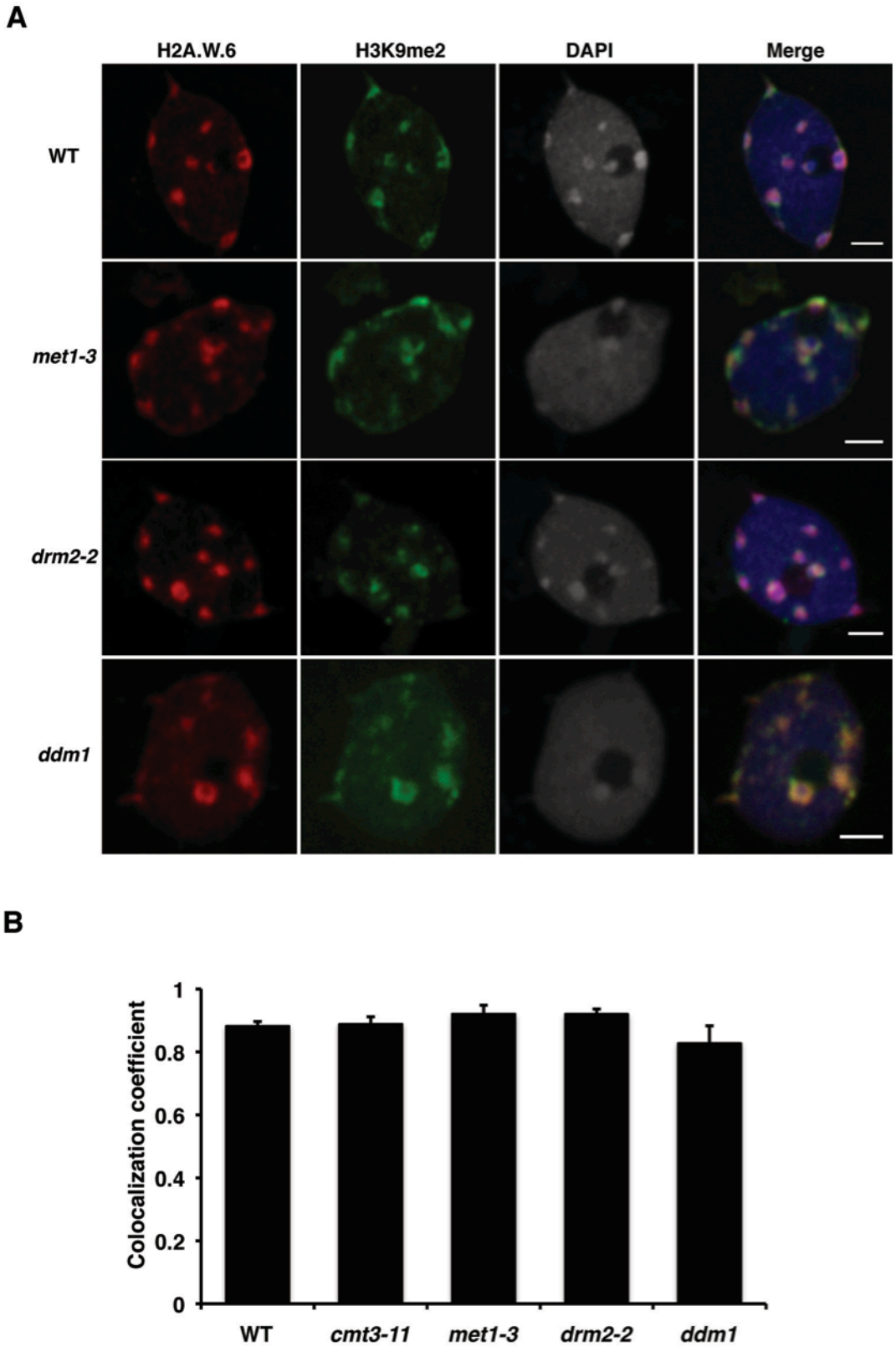

Figure 4.4. Recruitment of H2A.W.6 to chromatin is independent of DNA methylation

(A) Confocal sections of WT, *met1-3, drm2-2*, and *ddm1* leaf nuclei immunostained for H3K9me2 and H2A.W.6 with DAPI counterstain. H2A.W.6 is co-localized with H3K9me2 in all the genetic backgrounds. Scale bars represent 2 μm. (B) Quantification of H2A.W.6 and H3K9me2 co-localization in WT, *cmt3-11*, *met1-3, drm2-2*, and *ddm1* using Imaris software (Bitplane). The graph is plotted using Pearson's coefficient in ROI volume values on y-axis. Error bars represent standard deviation based on a minimum of three independent datasets.

4.4 Loss of H2A.W causes heterochromatin decondensation

To study the function of H2A.W *in vivo*, I obtained knock-out lines for all three genes encoding H2A.W (Figure 4.5A-B). In single mutants I did not observe any obvious visible phenotypes. In *h2a.w.6 h2a.w.7* or *h2a.w.6 h2a.w.12* double mutants significant growth defects were observed, and in *h2a.w.6 h2a.w.7 h2a.w.12* (*h2a.w*) triple mutants, plant growth was severely affected (Figure 4.5C). The strong effect of the loss of function allele *h2a.w.6* is mostly due to the predominant level of expression of *HTA6* among the three H2A.W encoding genes (Figure 4.1A). Triple mutant plants failed to flower and did not give rise to the next generation, suggesting that H2A.W is essential. Since H2A.W is specifically localized to heterochromatin, I analyzed heterochromatin structure in nuclei of *h2a.w* knock-out plants. In the double and triple mutants, a significant proportion of nuclei showed variable degrees of decondensed constitutive heterochromatin, marked by DAPI and still marked by H3K9me2 (Figures 4.6A-B). These results suggest that H2A.W is required for heterochromatin condensation.

Since heterochromatin is disorganized in *h2a.w* mutant, I analyzed the localization of the conserved heterochromatic mark H3K9me2 in *h2a.w* triple mutants. In this genetic background, H3K9me2 was still localized to DAPI rich regions even in nuclei with dispersed heterochromatic chromocenters, suggesting that H3K9me2 deposition does not require H2A.W (Figure 4.7A). Together with the persistent localization of H2A.W in absence of H3K9me2 we conclude that the deposition of H2A.W and H3K9me2 depend on mutually independent mechanisms.

The co-localization of H2A.W and H3K9me2 suggests that they might be working for a common function. In the H3K9 methyltransferase mutant *kyp*, H3K9me2 is reduced but

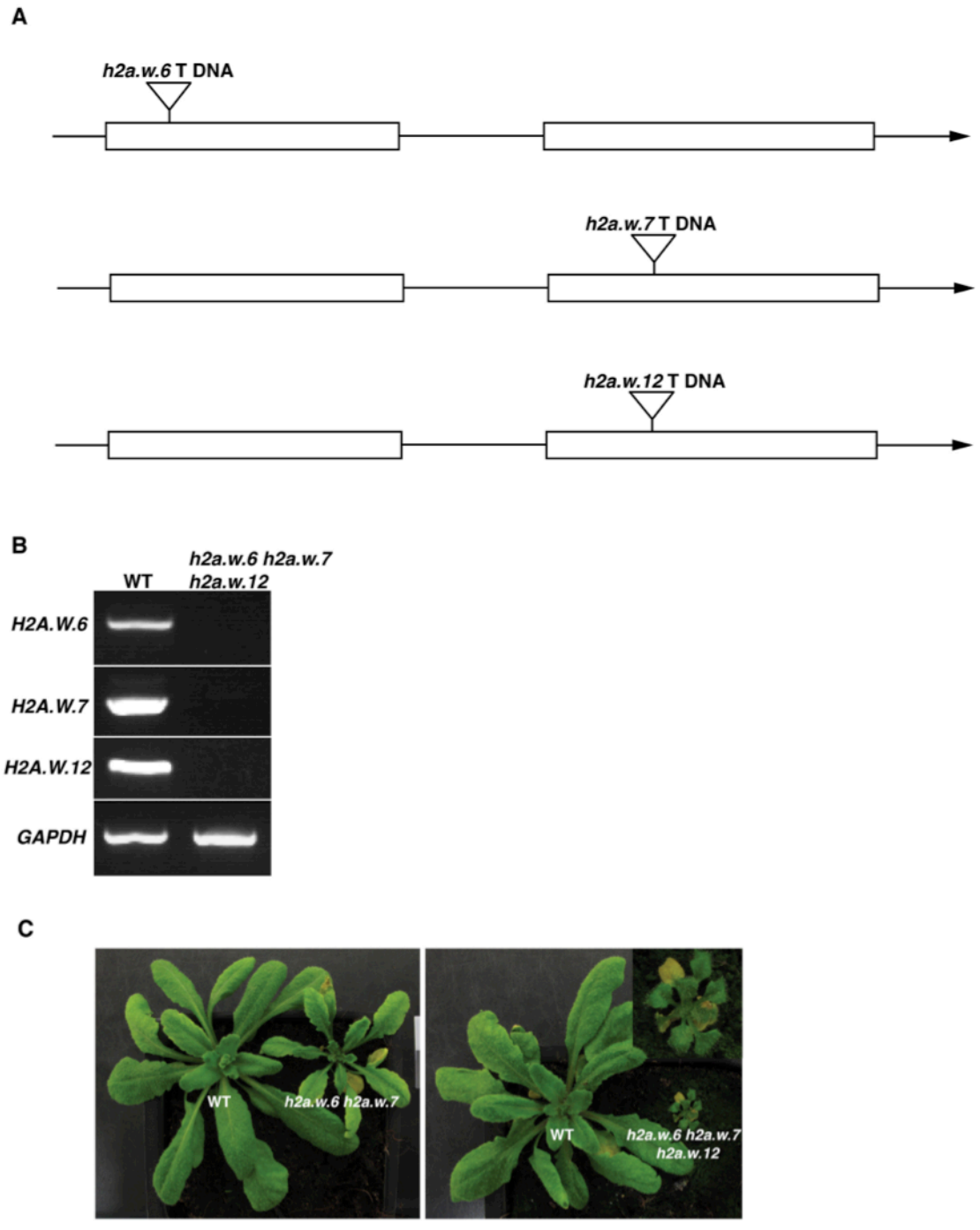

Figure 4.5. Loss of H2A.W causes severe growth phenotypes
(A) Schematic diagram showing the position of the T-DNA in *H2A.W* gene locus, Open boxes depict the exons in gene loci. (B) Semi quantitative RT-PCR of *H2A.W.6*, *H2A.W.7*, and *H2A.W.12* in *h2a.w.6 h2a.w.7 h2a.w.12 triple mutant* and WT. *GAPDH* was used as an endogenous control. (C) Morphological phenotypes observed in *h2a.w* mutants. Triple mutant is highlighted in inset.

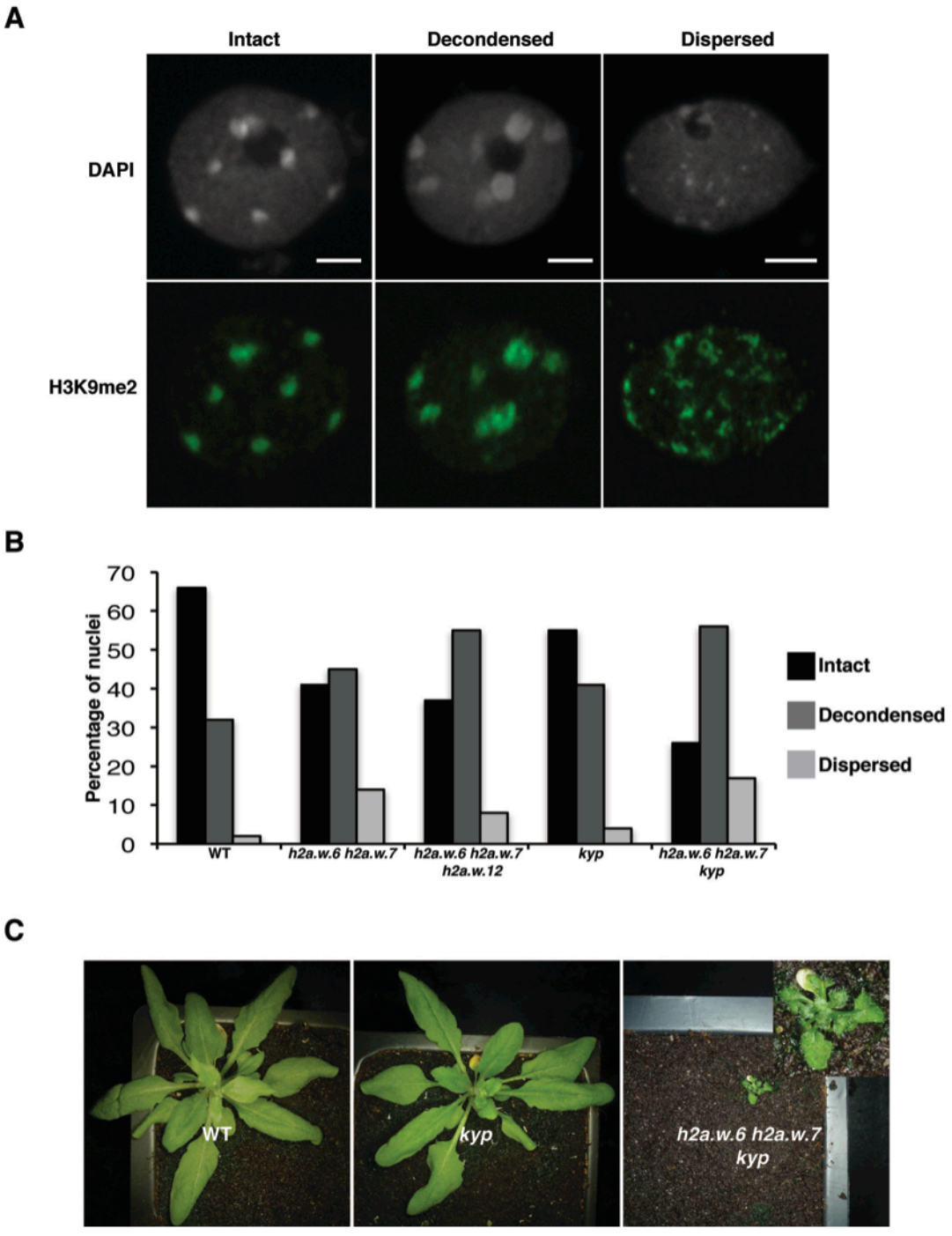

Figure 4.6. Loss of H2A.W causes decondensation of heterochromatin
(A) Decondensed chromocentres observed in leaf nuclei stained with DAPI. Scale bars represent 2 μm. (B) The bar chart represents percentages of nuclei belonging to each class of nuclei based on the degree of decondensation of chromocentres observed in DAPI staining. Where n=300 nuclei were observed for each sample. (C) Morphological phenotype observed in *h2a.w.6 h2a.w.7 kyp* mutant. Triple mutant is highlighted in inset.

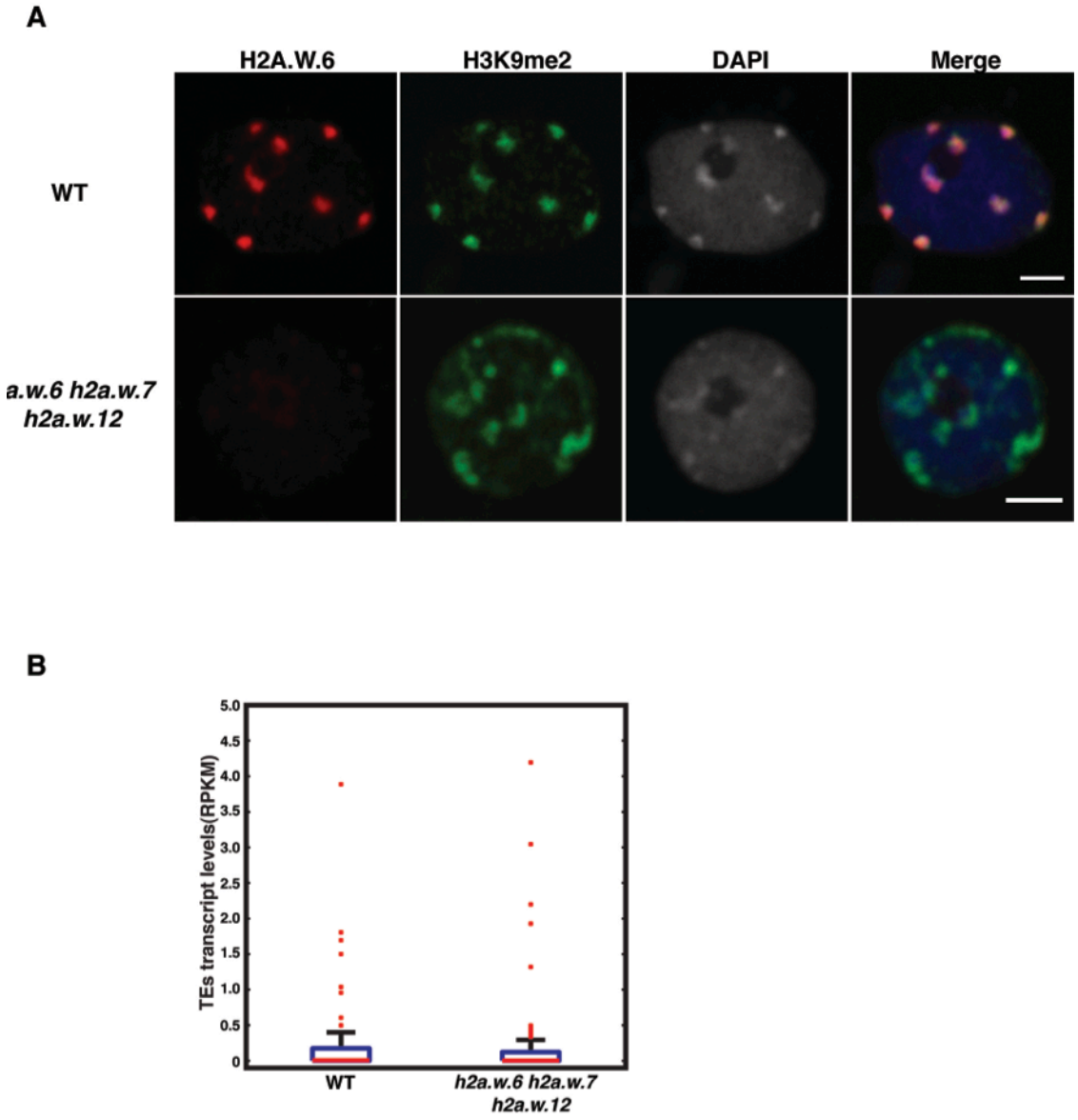

Figure 4.7. H3K9me2 localization is not dependent on H2A.W and in *h2a.w* mutant transposons expression is normal
(A) WT and *h2a.w.6 h2a.w.7 h2a.w.12* leaf nucleus immunostained for H3K9me2 and H2A.W.6 with DAPI counterstain. We did not detect H2A.W.6 protein in *h2a.w* triple mutant chromatin. Confocal sections, the scale bar represents 2 μm (B) Box plots showing the expression levels of TEs. TEs with 4 fold higher H2A.W.6 ChIP-seq read density compared to H3 and with a non-zero RPKM value for either WT or *h2a.w.6 h2a.w.7 h2a.w.12* triple mutant were included in the analysis.

does not show significant effects on plant growth (Jackson et al., 2002). Simultaneous reduction in dosage of *KYP* and H2.A.W in *h2a.w.6, h2a.w.7, kyp* plants showed severe growth defects compared to growth defects observed in *h2a.w.6, h2a.w.7* (Figures 4.6C and 4.5C) mutant plants. Similarly, I also observed an increase in the proportion of nuclei with decondensed heterochromatin (Figure 4.6B). Altogether, these results indicate that H2A.W and KYP participate in parallel pathways controlling heterochromatin condensation.

4.5 H2A.W represses transposable elements in synergy with CMT3 mediated CHG methylation

Decondensation of heterochromatin in other mutants such as *met1* has been previously shown to be associated with increased transcription of TEs and genes in the proximity of heterochromatic loci (Law and Jacobsen, 2010). However, RNA sequencing (RNA-seq) analyses of the *h2a.w* triple mutant did not show a substantial impact on expression of TEs, which are enriched in H2A.W in WT background (Figure 4.7B). This suggests that H2A.W plays a predominant role in condensation of heterochromatin within the nucleus, and that in the absence of H2A.W, other silencing pathways might play compensatory roles in maintaining the transcriptional repressive status of TEs. Indeed, DNA methylation analysis by whole genome bisulfite sequencing (BS-seq) showed a significant increase in CHG methylation over TEs and regions enriched in H3K9me2, while DNA methylation over gene bodies was not affected (Figure 4.8A). A notable up-regulation of the expression of the primary CHG DNA methyltransferase CMT3 (Figure 4.8B) likely accounts for the increase in CHG methylation. DNA methylation at CG and CHH sites were not significantly affected (Figure 4.8A), outlining the strong association of H2A.W with CHG methylation. It is thus likely that increased CHG methylation may compensate for the lack of H2A.W, resulting in maintenance of the transcriptional repression of TEs. To test this hypothesis, I combined the null mutant alleles, *h2a.w.6* and *h2a.w.7,* with *cmt3-11*. Combination of null alleles *cmt3-11*, *h2a.w.6*, *and h2a.w.7* led to male and female gamete abortion (Figure 4.8C) and I was unable to obtain any triple homozygous plant out of 704 progeny segregating from the self fertilized triple heterozygous mutant, indicating that CMT3 and H2A.W pathways are synergistically

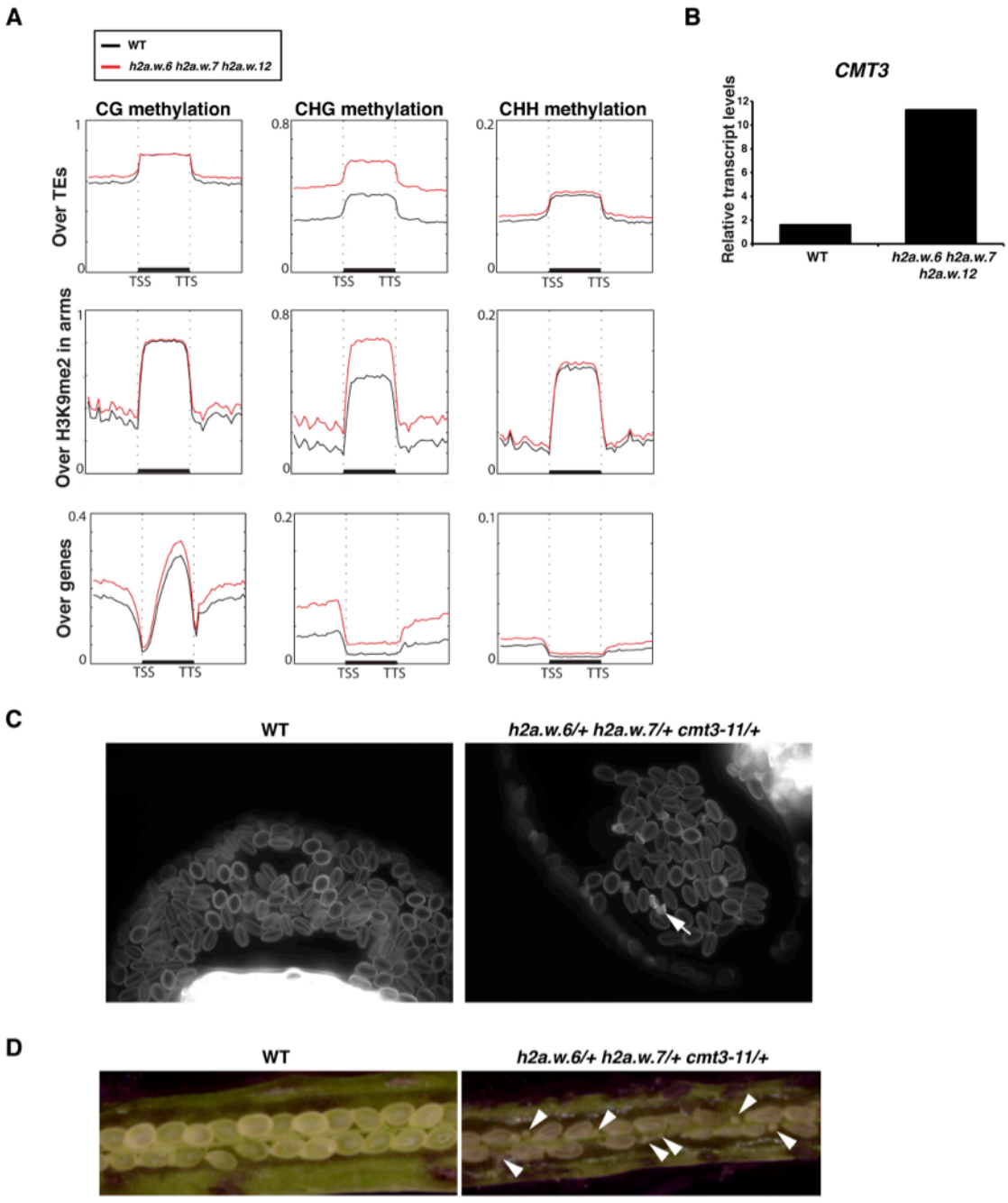

Figure 4.8. H2A.W and CMT3 genetically interact to repress heterochromatic elements (A) Metaplot analysis from whole genome bisulfite sequencing showing different types of methylation levels in WT and *h2a.w* triple mutant over major genomic features. Graphs were plotted as in Figure 3.5 (B) Expression level of *CMT3* in *h2a.w* triple mutant in comparison to WT. Expression values were obtained from RNA-seq data. (C) Abortion of pollen in anther of *h2a.w.6/+ h2a.w.7/+ cmt3-11/+* triple mutant compared with WT. (D) Abortion of ovule development in *h2a.w.6/+ h2a.w.7/+ cmt3-11/+* triple mutant compared with WT.

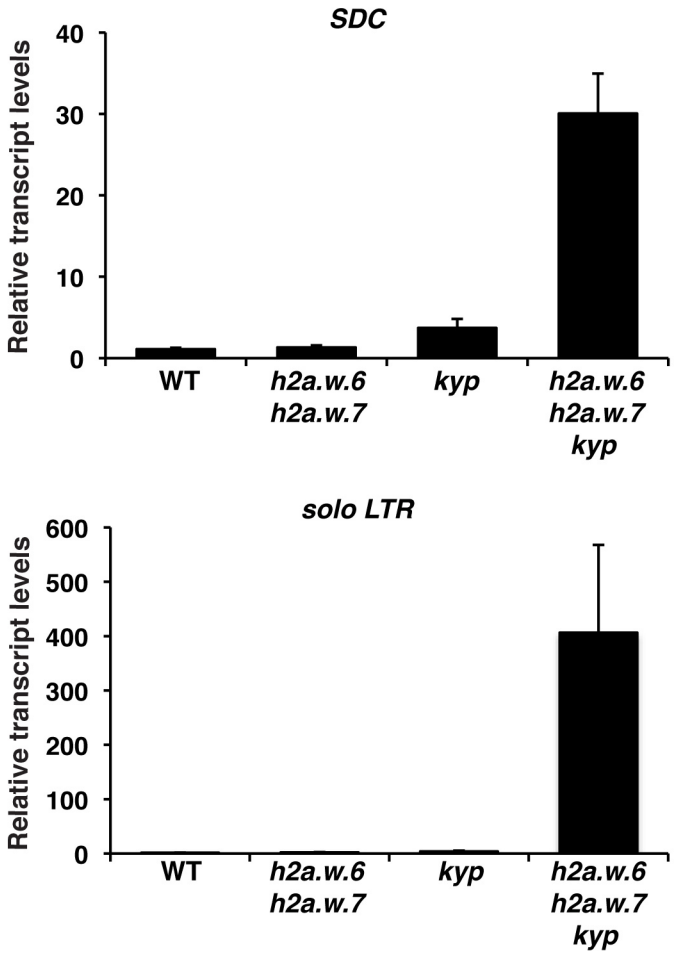

Figure 4.9. H2A.W represses the transcription of heterochromatic elements in synergy with H3K9me2

(A) Expression levels of TEs in *h2a.w.6 h2a.w.7 kyp* mutant compared with WT. Quantitative real-time RT-PCR was performed for *SDC* and *solo LTR* and the data was normalized to *ACTIN7*. Error bars represent standard deviation based on three independent biological replicates.

required for viability. CMT3 methylates heterochromatic DNA by binding to H3K9me2 (Du et al., Cell 2012), and consistently, CHG methylation is reduced in *kyp* mutants (Jackson et al., Nature 2002, Stroud et al., Cell 2013). Accordingly, the synergistic action of H2A.W and KYP was also observed on expression of TEs and the gene *SDC* that is regulated in a manner similar to TEs (Henderson and Jacobsen, 2008), which was dramatically increased in the *h2a.w.6 h2a.w.7 kyp* triple mutant (Figure 4.9). This suggests that genetic pathways of H2A.W and CHG methylation interact to ensure heterochromatin integrity.

4.6 H2A.W promotes chromatin condensation through its conserved C terminal motif *in vitro*

It is clear from loss function analysis that H2A.W is required for chromatin condensation and heterochromatin silencing. To further understand the potential mechanism of H2A.W in chromatin condensation, in collaboration with Karoline Luger lab at Colorado State University, we have tested whether H2A.W is able to cause condensation of nucleosome arrays *in vitro*. Chromatin condensation *in vitro* is composed of two independent structural transitions. Nucleosomes undergo short-range interactions with neighboring nucleosomes to form locally folded chromatin fibers. In addition, chromatin fibers self-associate to form supramolecular oligomers that share many of the properties of large-scale structure in condensed chromosomes. These two reversible structural transitions rely on different types of nucleosome-nucleosome interactions (Hansen, 2002), and can be studied in a model system by analysing the magnesium-dependent solution-state properties of defined 12-mer nucleosomal arrays *in vitro*. To understand the mechanism of H2A.W in chromatin condensation *in vitro*, we tested whether H2A.W affects self-association or folding. We assembled nucleosome arrays with *Arabidopsis* H2A.1, H2A.W.6, or a truncated version of H2A.W.6 (H2A.W-CT) that lacks an extended tail containing the SPKK motif (Figure 4.10). Because folding and self-association depend on the number of nucleosomes per array, the three types of arrays were closely matched for nucleosome saturation (Figure 4.10A). We demonstrate that all contain on average 11-12 nucleosomes per DNA molecule, i.e. every nucleosome positioning sequence repeat is occupied (Figures 4.10A-B). We also verified equimolar distribution of all histone

components by SDS-PAGE (Figure 4.10C). We then determined the extent of self-association as a function of salt concentration, and found

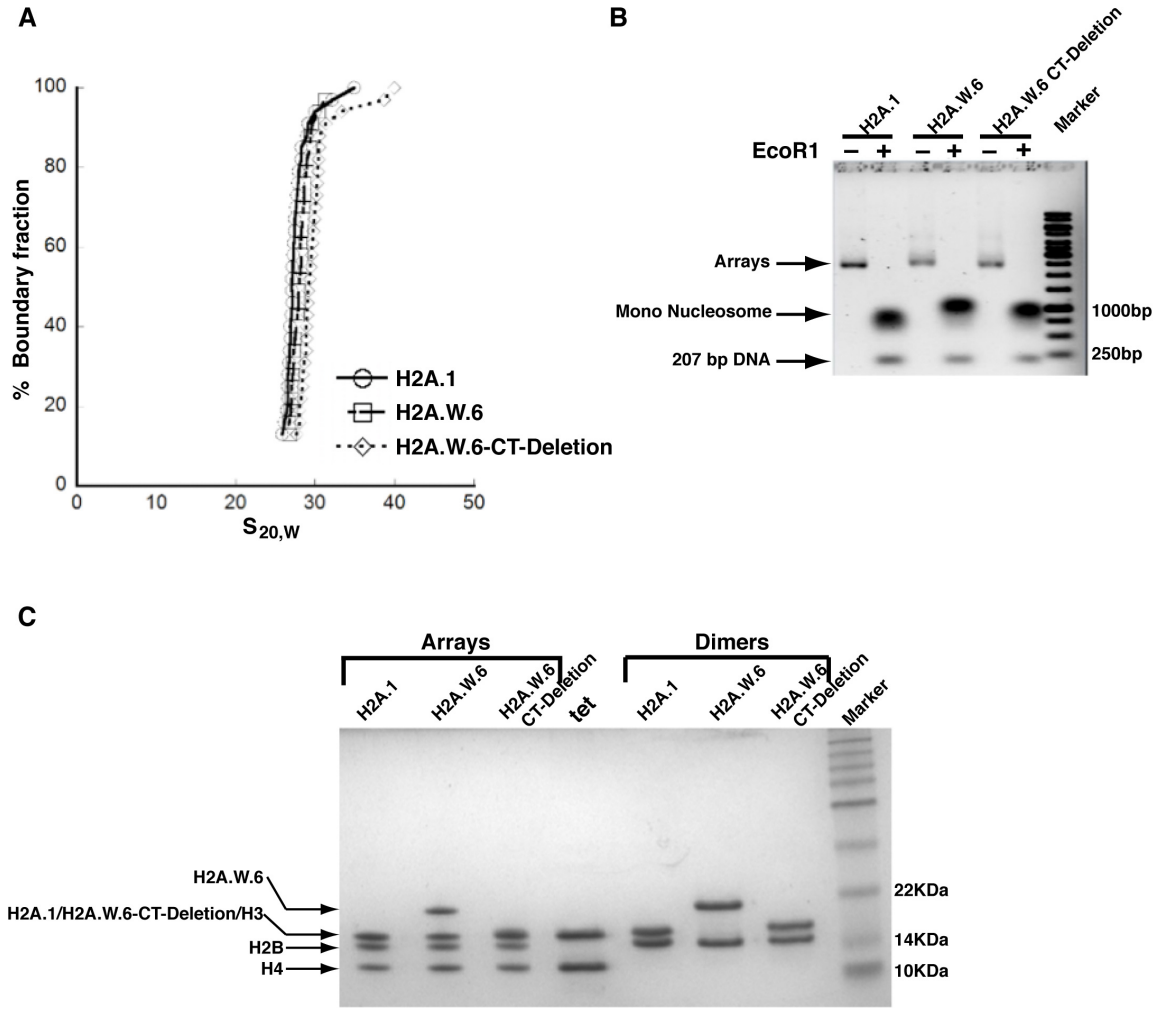

Figure 4.10. Quality control of nucleosome arrays
(A) SV-AUC profile for 207-12 arrays assembled with H2A.1, H2A.W.6, and H2A.W.6-CT deletion indicates similar (and near-complete) saturation. (B) EcoRI digestion of nucleosome arrays to assay the degree of saturation of each array. Each array exhibits similar amounts of free 207 bp DNA confirming similar saturation of the arrays. (C) 1-2 µg of each array from (A) was run on a 15% SDS-PAGE to confirm equimolar distribution of histones. Tet indicates $(H3\text{-}H4)_2$ tetramer. Note that H2A.1 and H2A.W.6-CT deletion runs in the same position as H3 on the gel.

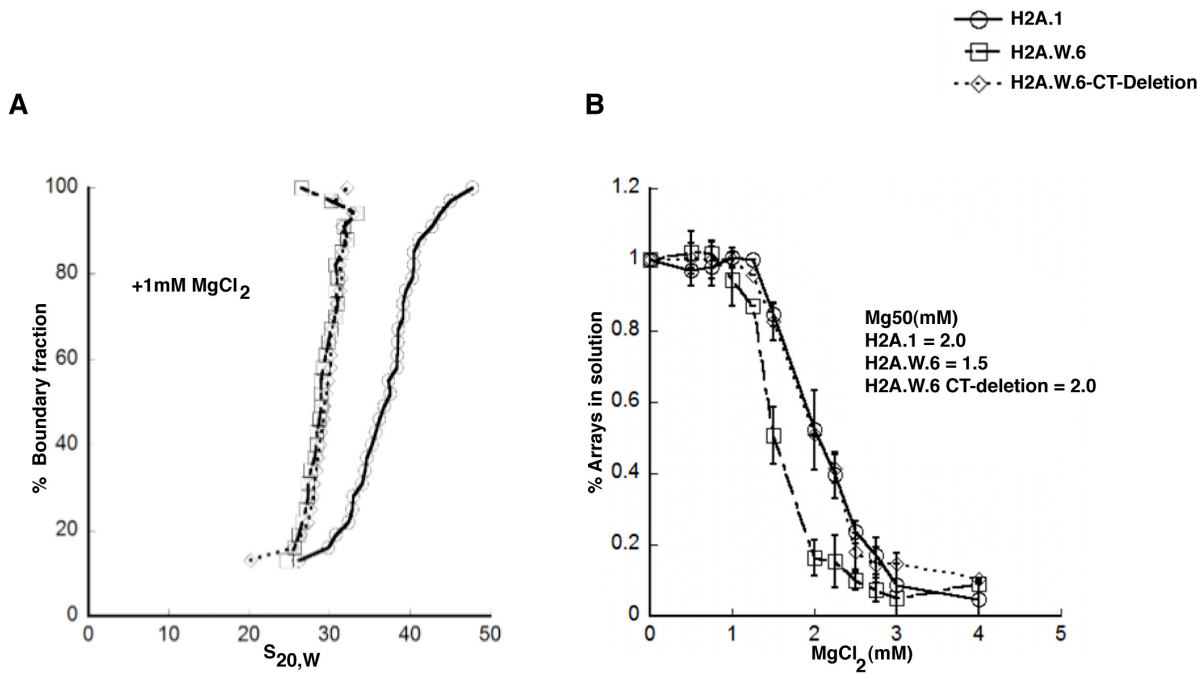

Figure 4.11. H2A.W causes chromatin condensation through chromatin fiber-fiber interactions promoted by its conserved C terminal tail *in vitro***.** (A) The arrays were subjected to folding assay in the presence of 1 mM $MgCl_2$. The H2A.1 containing control arrays show moderate folding (mid-point S ~40S), whereas both H2A.W.6 and H2A.W.6-CT deletion containing arrays do not undergo folding under these conditions. (B) Magnesium chloride-induced self-association assay showing that H2A.W.6 promotes inter-nucleosome array interactions. Arrays were incubated at the indicated concentrations of $MgCl_2$, and the concentration of Mg^{2+} at which 50 % of the arrays have precipitated (Mg_{50}) is listed in the figure (inset). Error bars are derived from two independent measurements (two preparations of reconstituted arrays).

that H2A.W.6 arrays engaged in oligomerization at significantly lower $MgCl_2$ concentrations than H2A.1 or H2A.W.6-CT (Figure 4.11A). This suggests that H2A.W.6 promotes long-range nucleosome-nucleosome interactions in chromosomes, and that the C-terminal tail of H2A.W.6 contributes to this behaviour. Intriguingly, both arrays with H2A.W.6 and H2A.W.6-CT exhibited no signs of folding at 1 mM $MgCl_2$, whereas arrays with canonical H2A.1 behave like other nucleosomal arrays and were moderately folded under these conditions (Figure 4.11B). Altogether, this suggests that H2A.W is able to enhance large-scale chromatin condensation by promoting long-range chromatin fiber-to-fiber interactions through its conserved C-terminal tail but is deficient in the short-range interactions that form the 30 nm fiber, suggesting a role in promoting highly condensed heterochromatin *in vivo*.

4.7 H2A.W promotes chromatin condensation *in vivo*

Since there was gain of DNA methylation upon loss of H2A.W in *h2a.w* triple mutants (Figure 4A), it was difficult to address directly whether H2A.W is sufficient to promote heterochromatin condensation in somatic cells. However, I observed that none of the genes encoding H2A.W was expressed in the central cell, one of the two female gametes of flowering plants (Figures 4.12A-C). The absence of H2A.W in this cell type correlated with the complete absence of condensed heterochromatic domains as indicated by the uniform distribution of H2A.X.5-RFP (Figure 4.13A), which marks both heterochromatin and euchromatin in somatic cells (Figures 3.4 and 3.5). In order to test whether the lack of H2A.W was responsible for the absence of condensed heterochromatin in the central cell nucleus, I ectopically expressed H2A.W.6 using the promoter of *FWA*, which is specifically active in the central cell and the endosperm (Kinoshita et al., 2004). The chromatin of the central cell expressing *pFWA-H2A.W.6-RFP* showed domains of varying intensity with distinct zones of higher condensation (Figure 4.13B). I could distinguish two types of organizations. A minor group of nuclei showed one or two strong patches of dense chromatin (Figure 4.13B). The most representative class of central cell nuclei showed several patches of dense chromatin assembled in periphery of the nucleus (Figure 4.13B), contrasting with the uniform wild-type pattern of H2A.X-RFP that marks both euchromatin and heterochromatin (Figure 4.13A). To test whether the assembled condensed domains are heterochromatic foci, I combined the *pFWA::H2A.W.6-RFP* transgenic line with *pCENH3::CENH3-GFP* line. In wild type CENH3-GFP is not detected in central cells (Ingouff et al., 2010), but in ~80% of central

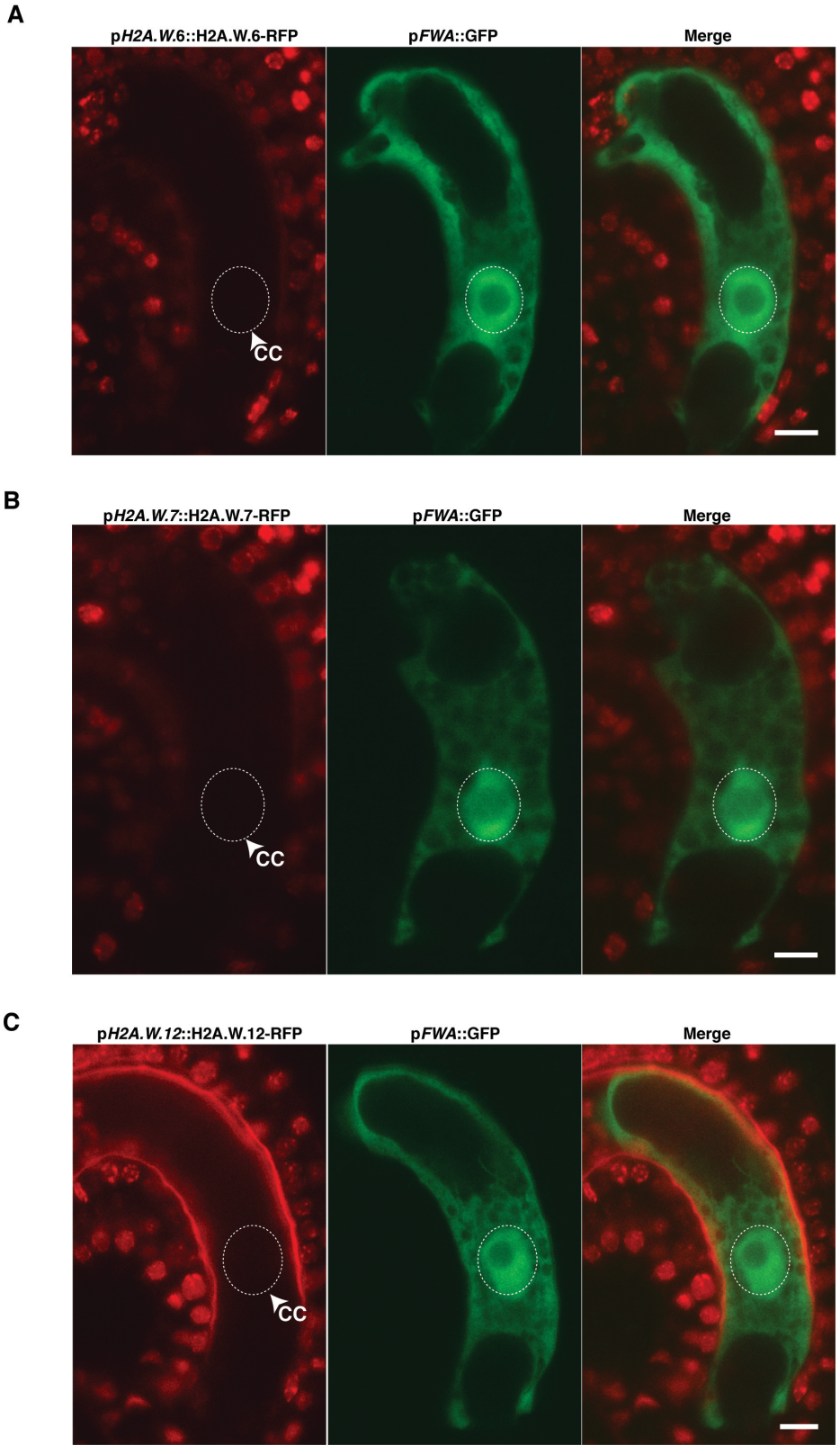

Figure 4.12. H2A.W is not expressed in the Arabidopsis female gametophyte central cell

(A) Absence of *pH2A.W.6::H2A.W.6-RFP* expression in the nucleus of the wild-type central cell (cc) marked by *pFWA::GFP* (dotted circle). (B) The expression of *pH2A.W.7-H2A.W.7-RFP* is not detected in the central cell nucleus (cc). (C) The expression of *pH2A.W.12-H2A.W.12-RFP* is not detected in the central cell nucleus (cc). In (A and B), the arrowhead and the dotted circle indicate the position of the central cell nucleus. Scale bars represent 5 μm.

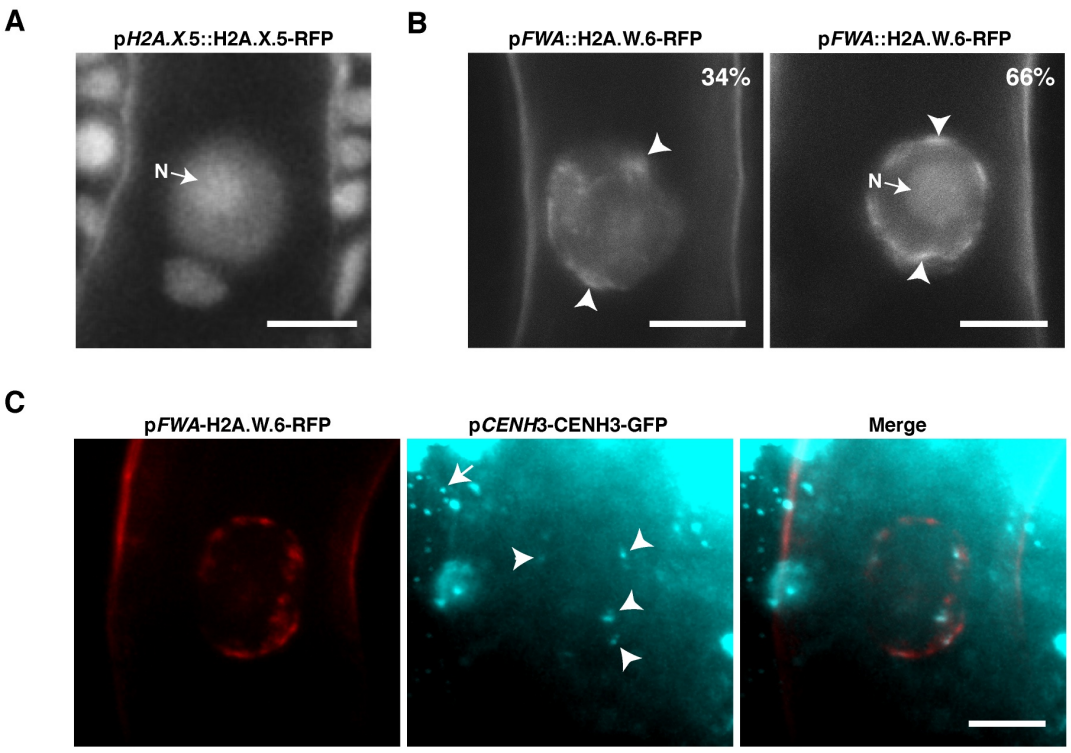

Figure 4.13. H2A.W is sufficient to cause chromatin condensation *in vivo*
(A) Uniform pattern of p*H2A.X.5*-H2A.X.5-RFP expressed in cc nucleus indicating the absence of defined chromocentres. N points the nucleolus. (B) Ectopic expression of *H2A.W.6-RFP* in cc nucleus showing condensed chromatin domains (arrowheads). Presented representative of three types of nuclei observed with their percentages in top right of each image. N points the nucleolus. (C) *In vivo* confocal section of a cc nucleus of transgenic plants expressing p*CENH3::CENH3-GFP* and *pFWA::H2A.W.6-RFP*. Arrowheads point the position of condensed chromatin domains and CENH3 position and arrow points the CENH3 localization in somatic cell. In **(**A to C), scale bars in confocal sections represent 5 μm.

cells expressing H2A.W.6-RFP, I consistently observed CENH3-GFP foci (Figure 4.13C), suggesting that condensed chromatin domains are indeed heterochromatin foci. These observations indicate that H2A.W is sufficient to cause heterochromatin condensation *in vivo*.

4.8 Discussion

Comprehensive genomic profiling showed that each H2A variant is characterized by a specific profile. In contrast with H2A, H2A.Z and H2A.W, the distribution of H2A.X the relatively ubiquitous, in prevention of a potential requirement of DNA repair (Downs et al., 2007; Soria et al., 2012). H2A marks uniformly gene bodies similar to that described for H3.1 (Stroud et al., 2012b; Wollmann et al., 2012). Both H3.1 and H2A transcription are highly coupled with S-phase (Talbert et al., 2012), suggesting that H2A and H3.1 are subjected to common dynamics related to the loading of new nucleosomes over replicated DNA. However, unlike H3.1, H2A enrichment over gene bodies positively correlates with transcription with a gradual replacement by H2A.Z in genes expressed at low levels, suggesting that transcription is coupled to an exchange of H2A and H2A.Z over gene bodies. I confirm that in highly expressed genes H2A.Z enrichment is confined to the 5' end of the gene bodies (Zilberman et al., 2008). This enrichment almost appears to be a mirror image of H3.3 distribution, confined to the 3' end of gene bodies (Stroud et al., 2012b; Wollmann et al., 2012). Hence, gene bodies boundaries appear to be defined by differential enrichment of H2A, H2A.Z and H3.3 variants. In sharp contrast with euchromatin, heterochromatin is depleted of H2A and H2A.Z and is occupied nearly exclusively by H2A.W suggesting that H2A variants restrict the deposition of each other over larger genomic domains defining euchromatin and heterochromatin. Therefore, combinatorial localizations of H2A and H3 variants define the main genomic features in *Arabidopsis*.

I show that H2A.W recruitment to chromatin is independent of the pathways that deposit heterochromatic marks H3K9me2 and DNA methylation. However, H2A.W acts together

with these pathways to maintain heterochromatin condensation into higher order domains and to silence TEs. The lethal genetic interaction observed in my study suggests that H2A.W and H3K9me2/CHG methylation are absolutely essential for plant life, although each pathway on their own is dispensable. The lethality observed in mutants deficient for both pathways likely results from a lack of heterochromatin integrity, which could impact cell division. In response to H2A.W depletion, I also observe a cross talk between these pathways as both CMT3 expression and CHG methylation increase in heterochromatin. This crosstalk is reminiscent of the cross-talk between H1 and DNA methylation and between the different pathways that control DNA methylation and H3K9me2, suggesting that sensing mechanisms report abnormal activity or condensation of heterochromatin(Mathieu et al., 2007; Zemach et al., 2013).

In vitro, nucleosomes containing H2A.W protect a longer stretch of DNA (162 bp) than nucleosomes containing H2A (146bp) to micrococcal nuclease digestion (Lindsey et al., 1991), suggesting that H2A.W prevents accessibility to DNA. Consistently, I find that H2A.W represses TE expression in synergistic action with H3K9me2. The acidic patch present in the histone core region of H2A affects the folding ability of nucleosome arrays (Kalashnikova et al., 2013). For example, euchromatic H2A.Z shows higher chromatin folding than H2A, as a consequence of its extended acidic patch (Fan et al., 2004). The acidic patch in H2A.W and canonical H2A are identical, and H2A.W shows a lower folding capacity than H2A, suggesting that other domains of H2A participate to chromatin folding. Major findings from my study support that H2A.W promotes chromatin condensation through long-range fiber-to-fiber interactions *in vitro* and *in vivo*. I propose that H2A.W is necessary and sufficient to organize heterochromatin into

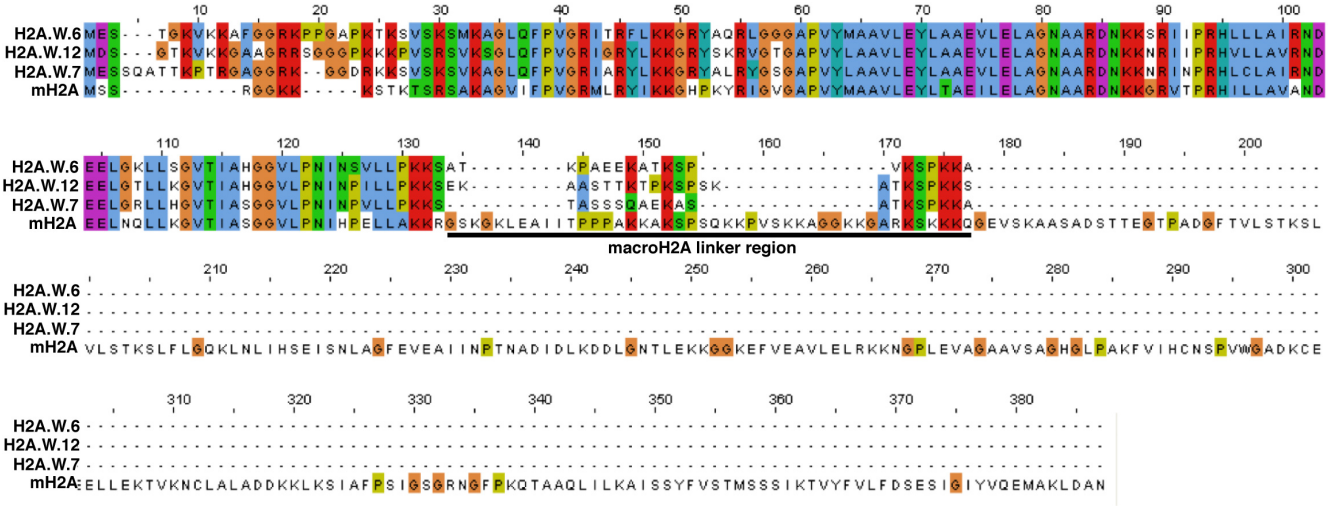

Figure 4.14. Conserved H2A.W C terminal like motifs are also present in linker region of mammalian specific macroH2A
Alignment of *Arabidopsis* H2A.W proteins with Human macroH2A generated with ClustalW2 program.

higher order functional domains such as chromocenters, which are essential for kinetochore assembly and mitosis (Bian and Belmont, 2012; Luger et al., 2012).

H2A.W specific chromatin condensation properties rely on the extended C-terminal domain that contains the conserved SPKK motif. This is consistent with previous *in vitro* studies of the function of the SPKK motif in other histone variants (Khadake and Rao, 1997). The SPKK motif binds preferentially to the minor groove of DNA at A/T rich sites (Churchill and Suzuki, 1989), which are generally enriched in satellite repeats of heterochromatin (Brutlag, 1980). In wheat the serine residue of the H2A.W SPKK motif is phosphorylated (Green et al., 1990). *In vitro* phosphorylation of SPKK reduces the affinity of wheat H2A.W to DNA (Green et al., 1993), suggesting that phosphorylation might regulate H2A.W-mediated heterochromatin condensation. It is plausible that during the S phase of DNA replication, phosphorylation of the SPKK motif by a cell cycle dependent kinase may increase the accessibility of heterochromatin to replication machinery.

During evolution, the SPKK motif not only became associated with H2A.W in seed plants but was also recruited to linker H1 and the sea urchin testis specific H2B (Suzuki, 1989), which are also associated with condensed chromatin (Buttinelli et al., 1999; Poccia and Green, 1992). Motifs similar to SPKK are also present in linker region of the vertebrate specific macroH2A (Figure 4.14), which is highly enriched in X-chromosome and is linked to repression (Costanzi and Pehrson, 1998a; Gamble et al., 2009; Gaspar-Maia et al., 2013). *In vitro*, the linker region of macroH2A which contains an SPKK like motif causes chromatin condensation (Muthurajan et al., 2011), and also stabilizes nucleosomes in the absence of macrodomain (Chakravarthy et al., 2012). However, how

macroH2A affects chromatin architecture *in vivo* remains unclear. Thus, my results uncover conserved aspects of histone variant functions that may not only aid in our understanding of plant heterochromatin, but also shed light on mechanisms controlling higher-order heterochromatin condensation in animals.

4.9 Conclusions

My comprehensive analysis of genome wide profiles of the four classes of H2A variants *in Arabidopsis* shows that H2A variants define the boundaries of the major genomic features, euchromatin, heterochromatin, gene bodies and their promoters. H2A variants act in concert with H3 variants. TSS are marked by H2A.Z while H3.3 marks TTS. I show that H2A variants occupancy on gene bodies is dependent on transcription as shown for H3.3.

I discovered the first histone variant that marks specifically heterochromatin in somatic cells. The heterochromatic H2A.W associates strictly with pericentromeric heterochromatin and H3K9me2 regions in chromosomal arms. I find that H2A.W and H3K9me2 are recruited to heterochromatin in a mutual independent manner. I find that H2A.W and the plant specific methyltransferase CMT3/KYP pathway, function cooperatively to repress the transcription of heterochromatic elements. Biochemical studies suggest that H2A.W promotes higher order chromatin condensation through chromatin fiber-to-fiber interactions through its conserved C terminal tail, which has the inherent capacity to bind DNA. Accordingly, I demonstrate that in vivo H2A.W is necessary and sufficient for heterochromatin condensation.

4.10 Future prospects

From this study, I find that the heterochromatic location of H2A.W is independent of other pathways that mark heterochromatin i.e H3K9me2 and DNA methylation. How H2A.W is recruited to heterochromatin remains a crucial point to solve that might reveal a novel mechanism. It is plausible that H2A.W is can be precisely recruited into the heterochromatin by using a dedicated chaperone, similar to the recruitment of CENH3 to centromeric heterochromatin by Scm3 and HJURP in yeast and humans respectively(Dunleavy et al., 2009; Foltz et al., 2009; Pidoux et al., 2009; Williams et al., 2009). The possibility of H2A.W dedicated chaperone can be investigated by immunoprecipitation of H2A.W associated protein complexes and these precipitated proteins can be identified by mass spectrometry. The identified proteins can be validated by genetic and biochemical studies.

It is also possible that H2A.W is removed from euchromatin by competition with H2A.Z chromatin remodelers, similar to H2A.X eviction(Downs et al., 2007; Soria et al., 2012). This is quite easy test; using immuno localization and ChIP sequencing we can check the H2A.W localization in H2A.Z chaperone mutants, such as *arp6* and *pie1*(March-Díaz and Reyes, 2009). The heterochromatin specific recruitment of H2A.W can also be achieved by the synchronized expression of H2A.W with the replication of heterochromatin during late S-phase. It is well known that heterochromatin is replicated towards the end of S-phase(Wallace and Orr-Weaver, 2005). It is also possible to envisage that a mechanism related to the cell cycle orchestrates a succession of deposition of H2A variants. The possibility of cell cycle regulated recruitment can be addressed by miss expression of H2A.W using cell cycle regulated promoters, which are active during different stages of

S-phase. The fact that H2A and H3 variants define functional domains in chromatin also alludes to the idea that a potential competition for occupancy structures chromatin organization. Thus, it will be interesting to test the chromatin status and H2A variants localization in the absence of each category of H2A and H3 variant.

The correlation between H2A variants enrichment on gene bodies and their transcription levels is intriguing. It recalls the link observed with H3.3 enrichment at TSS(Stroud et al., 2012b; Wollmann et al., 2012). It is not clear whether H2A and H3.3 variants control transcriptional activity or if transcription regulate the localization of H2A variants. This can be addressed by investigating the relation between transcriptional miss regulation of genes and changes in the localization of H2A variants in mutants of different H2A and H3 variants.

As discussed in the introduction, functions of histones also modulated by post translation modifications. H2As are target sites for phosphorylation and ubiquitination. Except for preliminary indication of phosphorylation of wheat H2A.W (Green et al., 1990), posttranslational modifications on H2As are not well studied in plants. Making use of biochemical approaches and latest mass spectrometry techniques, we can find out the modifications on H2As. By altering or systematic analysis of the modification sites, we can address the functions of those modifications and their effects chromatin organization.

Condensation of the chromosomes is an essential step during cell division and also plays a critical role in homologous recombination of the sister chromatids during the meiosis. Intriguingly, recent studies in *Arabidopsis* found that DNA methylation abnormalities, which also show effects on heterochromatin condensation, have severe effects on genetic recombination(Colomé-Tatché et al., 2012; Mirouze et al., 2012; Yelina et al., 2012). It

would thus be interesting to study the role of H2A.W mediated chromatin condensation in homologous recombination during the meiosis and in gametes segregation.

Chapter 5: References

Albig, W., Drabent, B., Kunz, J., Kalff-Suske, M., Grzeschik, K., and Doenecke, D. (1993). All known human H1 histone genes except the H1(0) gene are clustered on chromosome 6. Genomics *16*, 649 - 654.

Albig, W., Kioschis, P., Poustka, A., Meergans, K., and Doenecke, D. (1997). Human histone gene organization: nonregular arrangement within a large cluster. Genomics *40*, 314 - 322.

Bao, Y., Konesky, K., Park, Y., Rosu, S., Dyer, P., Rangasamy, D., Tremethick, D., Laybourn, P., and Luger, K. (2004). Nucleosomes containing the histone variant H2A.Bbd organize only 118 base pairs of DNA. EMBO J *23*, 3314 - 3324.

Bernatavichute, Y.V., Zhang, X., Cokus, S., Pellegrini, M., and Jacobsen, S.E. (2008). Genome-wide association of histone H3 lysine nine methylation with CHG DNA methylation in Arabidopsis thaliana. PLoS ONE *3*, e3156.

Bian, Q., and Belmont, A.S. (2012). Revisiting higher-order and large-scale chromatin organization. Current Opinion in Cell Biology *24*, 359-366.

Bönisch, C., and Hake, S.B. (2012). Histone H2A variants in nucleosomes and chromatin: more or less stable? Nucleic Acids Research *40*, 10719-10741.

Bönisch, C., Schneider, K., Pünzeler, S., Wiedemann, S.M., Bielmeier, C., Bocola, M., Eberl, H.C., Kuegel, W., Neumann, J., Kremmer, E.*, et al.* (2012). H2A.Z.2.2 is an alternatively spliced histone H2A.Z variant that causes severe nucleosome destabilization. Nucleic Acids Research *40*, 5951-5964.

Bruce, K., Myers, F.A., Mantouvalou, E., Lefevre, P., Greaves, I., Bonifer, C., Tremethick, D.J., Thorne, A.W., and Crane-Robinson, C. (2005). The replacement histone H2A.Z in a hyperacetylated form is a feature of active genes in the chicken. Nucleic Acids Research *33*, 5633-5639.

Brutlag, D.L. (1980). Molecular Arrangement and Evolution of Heterochromatic DNA. Annual Review of Genetics *14*, 121-144.

Buttinelli, M., Panetta, G., Rhodes, D., and Travers, A. (1999). The role of histone H1 in chromatin condensation and transcriptional repression. Genetica *106*, 117-124.

Carr, A.M., Dorrington, S.M., Hindley, J., Phear, G.A., Aves, S.J., and Nurse, P. (1994). Analysis of a histone H2A variant from fission yeast: evidence for a role in chromosome stability. Molec Gen Genet *245*, 628-635.

CHAKRAVARTHY, S., BAO, Y., ROBERTS, V.A., TREMETHICK, D., and LUGER, K. (2004). Structural Characterization of Histone H2A Variants. Cold Spring Harbor Symposia on Quantitative Biology *69*, 227-234.

Chakravarthy, S., Patel, A., and Bowman, G.D. (2012). The basic linker of macroH2A stabilizes DNA at the entry/exit site of the nucleosome. Nucleic Acids Research *40*, 8285-8295.

Checchi, P.M., and Engebrecht, J. (2011). Heteromorphic sex chromosomes: Navigating meiosis without a homologous partner. Molecular Reproduction and Development *78*, 623-632.

Churchill, M.E.A., and Suzuki, M. (1989). SPKK MOTIFS PREFER TO BIND TO DNA AT A/T-RICH SITES. Embo Journal *8*, 4189-4195.

Coleman-Derr, D., and Zilberman, D. (2012). Deposition of Histone Variant H2A.Z within Gene Bodies Regulates Responsive Genes. PLoS Genet *8*, e1002988.

Colomé-Tatché, M., Cortijo, S., Wardenaar, R., Morgado, L., Lahouze, B., Sarazin, A., Etcheverry, M., Martin, A., Feng, S., Duvernois-Berthet, E.*, et al.* (2012). Features of the Arabidopsis recombination landscape resulting from the combined loss of sequence variation and DNA methylation. Proceedings of the National Academy of Sciences *109*, 16240-16245.

Cook, P.J., Ju, B.G., Telese, F., Wang, X., Glass, C.K., and Rosenfeld, M.G. (2009). Tyrosine dephosphorylation of H2AX modulates apoptosis and survival decisions. Nature *458*, 591-596.

Costanzi, C., and Pehrson, J. (1998a). Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. Nature *393*, 599 - 601.

Costanzi, C., and Pehrson, J.R. (1998b). Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. Nature *393*, 599-601.

Creyghton, M.P., Markoulaki, S., Levine, S.S., Hanna, J., Lodato, M.A., Sha, K., Young, R.A., Jaenisch, R., and Boyer, L.A. (2008). H2AZ Is Enriched at Polycomb Complex Target Genes in ES Cells and Is Necessary for Lineage Commitment. Cell *135*, 649-661.

Csankovszki, G., Panning, B., Bates, B., Pehrson, J., and Jaenisch, R. (1999). Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X inactivation. Nat Genet *22*, 323 - 324.

Deal, R.B., Topp, C.N., McKinney, E.C., and Meagher, R.B. (2007). Repression of Flowering in Arabidopsis Requires Activation of FLOWERING LOCUS C Expression by the Histone Variant H2A.Z. The Plant Cell Online *19*, 74-83.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.-F., Guindon, S., Lefort, V., Lescot, M.*, et al.* (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Research *36*, W465-W469.

Downs, J.A., Nussenzweig, M.C., and Nussenzweig, A. (2007). Chromatin dynamics and the preservation of genetic information. Nature *447*, 951-958.

Doyen, C., An, W., Angelov, D., Bondarenko, V., Mietton, F., Studitsky, V., Hamiche, A., Roeder, R., Bouvet, P., and Dimitrov, S. (2006). Mechanism of polymerase II transcription repression by the histone variant macroH2A. Mol Cell Biol *26*, 1156 - 1164.

Du, J., Zhong, X., Bernatavichute, Y.V., Stroud, H., Feng, S., Caro, E., Vashisht, A.A., Terragni, J., Chin, H.G., Tu, A.*, et al.* (2012). Dual Binding of Chromomethylase Domains to H3K9me2-Containing Nucleosomes Directs DNA Methylation in Plants. Cell *151*, 167-180.

Dunleavy, E.M., Roche, D., Tagami, H., Lacoste, N., Ray-Gallet, D., Nakamura, Y., Daigo, Y., Nakatani, Y., and Almouzni-Pettinotti, G. (2009). HJURP Is a Cell-Cycle-Dependent Maintenance and Deposition Factor of CENP-A at Centromeres. Cell *137*, 485-497.

Earnshaw, W.C., and Tomkiel, J.E. (1992). Centromere and kinetochore structure. Current Opinion in Cell Biology *4*, 86-93.

Fan, J.Y., Rangasamy, D., Luger, K., and Tremethick, D.J. (2004). H2A.Z Alters the Nucleosome Surface to Promote HP1α-Mediated Chromatin Fiber Folding. Molecular cell *16*, 655-661.

Fang, Y., and Spector, D.L. (2005). Centromere Positioning and Dynamics in Living Arabidopsis Plants. Molecular Biology of the Cell *16*, 5710-5718.

Foltz, D.R., Jansen, L.E.T., Bailey, A.O., Yates Iii, J.R., Bassett, E.A., Wood, S., Black, B.E., and Cleveland, D.W. (2009). Centromere-Specific Assembly of CENP-A Nucleosomes Is Mediated by HJURP. Cell *137*, 472-484.

Fuchs, J., Demidov, D., Houben, A., and Schubert, I. (2006). Chromosomal histone modification patterns--from conservation to diversity. Trends in plant science *11*, 199-208.

Gamble, M., Frizzell, K., Yang, C., Krishnakumar, R., and Kraus, W. (2009). The histone variant macroH2A1 marks repressed autosomal chromatin, but protects a subset of its target genes from silencing. Genes Dev *24*, 21 - 32.

Gaspar-Maia, A., Qadeer, Z.A., Hasson, D., Ratnakumar, K., Adrian Leu, N., Leroy, G., Liu, S., Costanzi, C., Valle-Garcia, D., Schaniel, C.*, et al.* (2013). MacroH2A histone variants act as a barrier upon reprogramming towards pluripotency. Nat Commun *4*, 1565.

Goldberg, A.D., Banaszynski, L.A., Noh, K.-M., Lewis, P.W., Elsaesser, S.J., Stadler, S., Dewell, S., Law, M., Guo, X., Li, X., *et al.* (2010). Distinct Factors Control Histone Variant H3.3 Localization at Specific Genomic Regions. Cell *140*, 678-691.

Green, G.R., Gustavsen, L.C., and Poccia, D.L. (1990). Phosphorylation of Plant H2A Histones. Plant Physiology *93*, 1241-1245.

Green, G.R., Lee, H.J., and Poccia, D.L. (1993). Phosphorylation weakens DNA binding by peptides containing multiple "SPKK" sequences. Journal of Biological Chemistry *268*, 11247-11255.

Grewal, S.I., and Elgin, S.C. (2007). Transcription and RNA interference in the formation of heterochromatin. Nature *447*, 399-406.

Grosberg, A.Y. (2012). How two meters of DNA fit into a cell nucleus: Polymer models with topological constraints and experimental data. Polym Sci Ser C *54*, 1-10.

Guillemette, B., Bataille, A.R., Gévry, N., Adam, M., Blanchette, M., Robert, F., and Gaudreau, L. (2005). Variant Histone H2A.Z Is Globally Localized to the Promoters of Inactive Yeast Genes and Regulates Nucleosome Positioning. PLoS Biol *3*, e384.

Haag, J.R., and Pikaard, C.S. (2011). Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. Nat Rev Mol Cell Biol *12*, 483-492.

Hammoud, S.S., Nix, D.A., Zhang, H., Purwar, J., Carrell, D.T., and Cairns, B.R. (2009). Distinctive chromatin in human sperm packages genes for embryo development. Nature *460*, 473-478.

Han, W., Li, X., and Fu, X. (2011). The macro domain protein family: Structure, functions, and their potential therapeutic implications. Mutation Research/Reviews in Mutation Research *727*, 86-103.

Hansen, J.C. (2002). CONFORMATIONAL DYNAMICS OF THE CHROMATIN FIBER IN SOLUTION: Determinants, Mechanisms, and Functions. Annual Review of Biophysics and Biomolecular Structure *31*, 361-392.

Hardy, S., Jacques, P.-É., Gévry, N., Forest, A., Fortin, M.-È., Laflamme, L., Gaudreau, L., and Robert, F. (2009). The Euchromatic and Heterochromatic Landscapes Are Shaped by Antagonizing Effects of Transcription on H2A.Z Deposition. PLoS Genet *5*, e1000687.

He, Y. (2012). Chromatin regulation of flowering. Trends in plant science *17*, 556-562.
Heitz, E. (1928). Das Heterochromatin der Moose. I Jahrb Wiss Botanik *69*, 762 - 818.

Henderson, I.R., and Jacobsen, S.E. (2008). Tandem repeats upstream of the Arabidopsis endogene SDC recruit non-CG DNA methylation and initiate siRNA spreading. Genes & Development *22*, 1597-1606.

Hu, G., Cui, K., Northrup, D., Liu, C., Wang, C., Tang, Q., Ge, K., Levens, D., Crane-Robinson, C., and Zhao, K. (2013). H2A.Z Facilitates Access of Active and Repressive Complexes to Chromatin in Embryonic Stem Cell Self-Renewal and Differentiation. Cell Stem Cell *12*, 180-192.

Ingouff, M., Rademacher, S., Holec, S., Šoljić, L., Xin, N., Readshaw, A., Foo, S.H., Lahouze, B., Sprunck, S., and Berger, F. (2010). Zygotic Resetting of the HISTONE 3 Variant Repertoire Participates in Epigenetic Reprogramming in Arabidopsis. Current Biology *20*, 2137-2143.

Ishibashi, T., Li, A., Eirin-Lopez, J., Zhao, M., Missiaen, K., Abbott, D., Meistrich, M., Hendzel, M., and Ausio, J. (2010). H2A.Bbd: an X-chromosome-encoded histone involved in mammalian spermiogenesis. Nucleic Acids Res *38*, 1780 - 1789.

Jackson, J.P., Lindroth, A.M., Cao, X., and Jacobsen, S.E. (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature *416*, 556-560.

Jacob, Y., Feng, S., LeBlanc, C.A., Bernatavichute, Y.V., Stroud, H., Cokus, S., Johnson, L.M., Pellegrini, M., Jacobsen, S.E., and Michaels, S.D. (2009). ATXR5 and ATXR6 are H3K27 monomethyltransferases required for chromatin structure and gene silencing. Nat Struct Mol Biol *16*, 763-768.

Jensen, K., Santisteban, M., Urekar, C., and Smith, M.M. (2011). Histone H2A.Z acid patch residues required for deposition and function. Mol Genet Genomics *285*, 287-296.
Jenuwein, T., and Allis, C.D. (2001). Translating the Histone Code. Science *293*, 1074-1080.

Jin, C., and Felsenfeld, G. (2007). Nucleosome stability mediated by histone variants H3.3 and H2A.Z. Genes & Development *21*, 1519-1529.

Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., and Felsenfeld, G. (2009). H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. Nat Genet *41*, 941-945.

Johnson, L.M., Law, J.A., Khattar, A., Henderson, I.R., and Jacobsen, S.E. (2008). SRA-domain proteins required for DRM2-mediated de novo DNA methylation. PLoS Genet *4*, e1000280.

Jullien, P.E., Susaki, D., Yelagandula, R., Higashiyama, T., and Berger, F.d.r. DNA Methylation Dynamics during Sexual Reproduction in Arabidopsis thaliana. Current Biology.

Kalashnikova, A.A., Porter-Goff, M.E., Muthurajan, U.M., Luger, K., and Hansen, J.C. (2013). The role of the nucleosome acidic patch in modulating higher order chromatin structure. Journal of The Royal Society Interface *10*.

Khadake, J.R., and Rao, M.R.S. (1997). Condensation of DNA and Chromatin by an SPKK-Containing Octapeptide Repeat Motif Present in the C-Terminus of Histone H1¸Ä†. Biochemistry *36*, 1041-1051.

Kinoshita, T., Miura, A., Choi, Y., Kinoshita, Y., Cao, X., Jacobsen, S.E., Fischer, R.L., and Kakutani, T. (2004). One-Way Control of FWA Imprinting in Arabidopsis Endosperm by DNA Methylation. Science *303*, 521-523.

Kitada, T., Schleker, T., Sperling, A.S., Xie, W., Gasser, S.M., and Grunstein, M. (2011). γH2A is a component of yeast heterochromatin required for telomere elongation. Cell Cycle *10*, 293-300.

Ku, M., Jaffe, J., Koche, R., Rheinbay, E., Endoh, M., Koseki, H., Carr, S., and Bernstein, B. (2012). H2A.Z landscapes and dual modifications in pluripotent and multipotent stem cells underlie complex genome regulatory functions. Genome Biology *13*, R85.

Kumar, S.V., and Wigge, P.A. (2010). H2A.Z-Containing Nucleosomes Mediate the Thermosensory Response in Arabidopsis. Cell *140*, 136-147.

Lantermann, A.B., Straub, T., Stralfors, A., Yuan, G.-C., Ekwall, K., and Korber, P. (2010). Schizosaccharomyces pombe genome-wide nucleosome mapping reveals positioning mechanisms distinct from those of Saccharomyces cerevisiae. Nat Struct Mol Biol *17*, 251-257.

Law, J.A., and Jacobsen, S.E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet *11*, 204-220.

Leach, T., Mazzeo, M., Chotkowski, H., Madigan, J., Wotring, M., and Glaser, R. (2000). Histone H2A.Z is widely but nonrandomly distributed in chromosomes of Drosophila melanogaster. J Biol Chem *275*, 23267 - 23272.

Li, A., Yu, Y., Lee, S.-C., Ishibashi, T., Lees-Miller, S.P., and Ausió, J. (2010). Phosphorylation of Histone H2A.X by DNA-dependent Protein Kinase Is Not Affected by Core Histone Acetylation, but It Alters Nucleosome Stability and Histone H1 Binding. Journal of Biological Chemistry *285*, 17778-17788.

Lindroth, A.M. (2001). Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. Science *292*, 2077-2080.

Lindsey, G., Orgeig, S., Thompson, P., Davies, N., and Maeder, D. (1991). Extended C-terminal tail of wheat histone H2A interacts with DNA of the "linker" region. J Mol Biol *218*, 805 - 813.

Lowary, P.T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. Journal of Molecular Biology *276*, 19-42.

Lu, P.Y.T., Lévesque, N., and Kobor, M.S. (2009). NuA4 and SWR1-C: two chromatin-modifying complexes with overlapping functions and componentsThis paper is one of a selection of papers published in this Special Issue, entitled 30th Annual International Asilomar Chromatin and Chromosomes Conference, and has undergone the Journal's usual peer review process. Biochemistry and Cell Biology *87*, 799-815.

Luger, K., Dechassa, M.L., and Tremethick, D.J. (2012). New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? Nat Rev Mol Cell Biol *13*, 436-447.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8[thinsp]A resolution. Nature *389*, 251-260.

Malagnac, F., Bartee, L., and Bender, J. (2002). An Arabidopsis SET domain protein required for maintenance but not establishment of DNA methylation. EMBO J *21*, 6842-6852.

Mannironi, C., Bonner, W.M., and Hatch, C.L. (1989). H2A.X. a histone isoprotein with a conserved C-terminal sequence, is encoded by a novel mRNA with both DNA replication type and polyA 3′ processing signals. Nucleic Acids Research *17*, 9113-9126.

March-Díaz, R., García-Domínguez, M., Lozano-Juste, J., León, J., Florencio, F.J., and Reyes, J.C. (2008). Histone H2A.Z and homologues of components of the SWR1 complex are required to control immunity in Arabidopsis. The Plant Journal *53*, 475-487.

March-Díaz, R., and Reyes, J.C. (2009). The Beauty of Being a Variant: H2A.Z and the SWR1 Complex in Plants. Molecular Plant *2*, 565-577.

Mathieu, O., Reinders, J., Čaikovski, M., Smathajitt, C., and Paszkowski, J. (2007). Transgenerational Stability of the Arabidopsis Epigenome Is Coordinated by CG Methylation. Cell *130*, 851-862.

Mavrich, T.N., Jiang, C., Ioshikhes, I.P., Li, X., Venters, B.J., Zanton, S.J., Tomsho, L.P., Qi, J., Glaser, R.L., Schuster, S.C.*, et al.* (2008). Nucleosome organization in the Drosophila genome. Nature *453*, 358-362.

Meneghini, M.D., Wu, M., and Madhani, H.D. (2003). Conserved Histone Variant H2A.Z Protects Euchromatin from the Ectopic Spread of Silent Heterochromatin. Cell *112*, 725-736.

Mietton, F., Sengupta, A.K., Molla, A., Picchi, G., Barral, S., Heliot, L., Grange, T., Wutz, A., and Dimitrov, S. (2009). Weak but Uniform Enrichment of the Histone Variant macroH2A1 along the Inactive X Chromosome. Molecular and Cellular Biology *29*, 150-156.

Millar, C.B. (2013). Organizing the genome with H2A histone variants. Biochemical Journal *449*, 567-579.

Mirouze, M., Lieberman-Lazarovich, M., Aversano, R., Bucher, E., Nicolet, J., Reinders, J., and Paszkowski, J. (2012). Loss of DNA methylation affects the recombination landscape in Arabidopsis. Proceedings of the National Academy of Sciences *109*, 5880-5885.

Mito, Y., Henikoff, J.G., and Henikoff, S. (2005). Genome-scale profiling of histone H3.3 replacement patterns. Nat Genet *37*, 1090-1097.

Mizuguchi, G., Shen, X., Landry, J., Wu, W.-H., Sen, S., and Wu, C. (2004). ATP-Driven Exchange of Histone H2AZ Variant Catalyzed by SWR1 Chromatin Remodeling Complex. Science *303*, 343-348.

Muthurajan, U.M., McBryant, S.J., Lu, X., Hansen, J.C., and Luger, K. (2011). The Linker Region of MacroH2A Promotes Self-association of Nucleosomal Arrays. Journal of Biological Chemistry *286*, 23852-23864.

Nashun, B., Yukawa, M., Liu, H., Akiyama, T., and Aoki, F. (2010). Changes in the nuclear deposition of histone H2A variants during pre-implantation development in mice. Development *137*, 3785-3794.

Park, Y.-J., Dyer, P.N., Tremethick, D.J., and Luger, K. (2004). A New Fluorescence Resonance Energy Transfer Approach Demonstrates That the Histone Variant H2AZ Stabilizes the Histone Octamer within the Nucleosome. Journal of Biological Chemistry *279*, 24274-24282.

Pasque, V., Gillich, A., Garrett, N., and Gurdon, J.B. (2011). Histone variant macroH2A confers resistance to nuclear reprogramming. EMBO J *30*, 2373-2387.

Pasque, V., Radzisheuskaya, A., Gillich, A., Halley-Stott, R.P., Panamarova, M., Zernicka-Goetz, M., Surani, M.A., and Silva, J.C.R. (2012). Histone variant macroH2A marks embryonic differentiation in vivo and acts as an epigenetic barrier to induced pluripotency. Journal of Cell Science *125*, 6094-6104.

Pehrson, J., and Fried, V. (1992). MacroH2A, a core histone containing a large nonhistone region. Science *257*, 1398 - 1400.

Pidoux, A.L., Choi, E.S., Abbott, J.K.R., Liu, X., Kagansky, A., Castillo, A.G., Hamilton, G.L., Richardson, W., Rappsilber, J., He, X*., et al.* (2009). Fission Yeast Scm3: A CENP-

A Receptor Required for Integrity of Subkinetochore Chromatin. Molecular cell *33*, 299-311.

Pinto, D., and Flaus, A. (2010). Structure and Function of Histone H2AX. In Genome Stability and Human Diseases, H.-P. Nasheuer, ed. (Springer Netherlands), pp. 55-78.

Poccia, D.L., and Green, G.R. (1992). Packaging and unpackaging the sea urchin sperm genome. Trends in Biochemical Sciences *17*, 223-227.

Prigent, C., and Dimitrov, S. (2003). Phosphorylation of serine 10 in histone H3, what for? Journal of Cell Science *116*, 3677-3685.

Raisner, R.M., Hartley, P.D., Meneghini, M.D., Bao, M.Z., Liu, C.L., Schreiber, S.L., Rando, O.J., and Madhani, H.D. (2005). Histone Variant H2A.Z Marks the 5′ Ends of Both Active and Inactive Genes in Euchromatin. Cell *123*, 233-248.

Rangasamy, D., Berven, L., Ridgway, P., and Tremethick, D.J. (2003). Pericentric heterochromatin becomes enriched with H2A.Z during early mammalian development. EMBO J *22*, 1599-1607.

Rangasamy, D., Greaves, I., and Tremethick, D.J. (2004). RNA interference demonstrates a novel role for H2A.Z in chromosome segregation. Nat Struct Mol Biol *11*, 650-655.

Ranjan, A., Mizuguchi, G., FitzGerald, Peter C., Wei, D., Wang, F., Huang, Y., Luk, E., Woodcock, Christopher L., and Wu, C. (2013). Nucleosome-free Region Dominates Histone Acetylation in Targeting SWR1 to Promoters for H2A.Z Replacement. Cell *154*, 1232-1245.

Ratnakumar, K., Duarte, L.F., LeRoy, G., Hasson, D., Smeets, D., Vardabasso, C., Bönisch, C., Zeng, T., Xiang, B., Zhang, D.Y*., et al.* (2012). ATRX-mediated chromatin association of histone variant macroH2A1 regulates α-globin expression. Genes & Development *26*, 433-438.

Ray−Gallet, D., and Almouzni, G. (2010). Nucleosome dynamics and histone variants. Essays in Biochemistry *48*, 75-87.

Sarma, K., and Reinberg, D. (2005). Histone variants meet their match. Nat Rev Mol Cell Biol *6*, 139-149.

Seo, J., Kim, S.C., Lee, H.-S., Kim, J.K., Shon, H.J., Salleh, N.L.M., Desai, K.V., Lee, J.H., Kang, E.-S., Kim, J.S*., et al.* (2012). Genome-wide profiles of H2AX and γ-H2AX differentiate endogenous and exogenous DNA damage hotspots in human cells. Nucleic Acids Research *40*, 5965-5974.

Smith, A.P., Jain, A., Deal, R.B., Nagarajan, V.K., Poling, M.D., Raghothama, K.G., and Meagher, R.B. (2010). Histone H2A.Z Regulates the Expression of Several Classes of

Phosphate Starvation Response Genes But Not as a Transcriptional Activator. Plant Physiology *152*, 217-225.

Soboleva, T.A., Nekrasov, M., Pahwa, A., Williams, R., Huttley, G.A., and Tremethick, D.J. (2012). A unique H2A histone variant occupies the transcriptional start site of active genes. Nat Struct Mol Biol *19*, 25-30.

Soria, G., Polo, Sophie E., and Almouzni, G. (2012). Prime, Repair, Restore: The Active Role of Chromatin in the DNA Damage Response. Molecular cell *46*, 722-734.

Stroud, H., Greenberg, M.V.C., Feng, S., Bernatavichute, Y.V., and Jacobsen, S.E. (2013). Comprehensive Analysis of Silencing Mutants Reveals Complex Regulation of the Arabidopsis Methylome. Cell *152*, 352-364.

Stroud, H., Hale, C.J., Feng, S., Caro, E., Jacob, Y., Michaels, S.D., and Jacobsen, S.E. (2012a). DNA Methyltransferases Are Required to Induce Heterochromatic Re-Replication in Arabidopsis. PLoS Genet *8*, e1002808.

Stroud, H., Otero, S., Desvoyes, B., Ramírez-Parra, E., Jacobsen, S.E., and Gutierrez, C. (2012b). Genome-wide analysis of histone H3.1 and H3.3 variants in Arabidopsis thaliana. Proceedings of the National Academy of Sciences *109*, 5370-5375.

Suto, R.K., Clarkson, M.J., Tremethick, D.J., and Luger, K. (2000). Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. Nat Struct Mol Biol *7*, 1121-1124.

Suzuki, M. (1989). SPKK, A NEW NUCLEIC ACID-BINDING UNIT OF PROTEIN FOUND IN HISTONE. Embo Journal *8*, 797-804.

Szilard, R.K., Jacques, P.-E., Laramee, L., Cheng, B., Galicia, S., Bataille, A.R., Yeung, M., Mendez, M., Bergeron, M., Robert, F.*, et al.* (2010). Systematic identification of fragile sites via genome-wide location analysis of [gamma]-H2AX. Nat Struct Mol Biol *17*, 299-305.

Talbert, P., Ahmad, K., Almouzni, G., Ausio, J., Berger, F., Bhalla, P., Bonner, W., Cande, W., Chadwick, B., Chan, S.W.*, et al.* (2012). A unified phylogeny-based nomenclature for histone variants. Epigenetics & Chromatin *5*, 7.

Talbert, P., and Henikoff, S. (2010). Histone variants-ancient wrap artists of the epigenome. Nat Rev Mol Cell Biol *11*, 264 - 275.

Talbert, P., Masuelli, R., Tyagi, A., Comai, L., and Henikoff, S. (2002). Centromeric localization and adaptive evolution of an Arabidopsis histone H3 variant. Plant Cell *14*, 1053 - 1066.

Tanasijevic, B., and Rasmussen, T.P. (2011). X Chromosome Inactivation and Differentiation Occur Readily in ES Cells Doubly-Deficient for MacroH2A1 and MacroH2A2. PLoS ONE *6*, e21512.

Tolstorukov, Michael Y., Goldman, Joseph A., Gilbert, C., Ogryzko, V., Kingston, Robert E., and Park, Peter J. (2012). Histone Variant H2A.Bbd Is Associated with Active Transcription and mRNA Processing in Human Cells. Molecular cell *47*, 596-607.

Turner, J.M.A. (2007). Meiotic sex chromosome inactivation. Development *134*, 1823-1831.

Valdés-Mora, F., Song, J.Z., Statham, A.L., Strbenac, D., Robinson, M.D., Nair, S.S., Patterson, K.I., Tremethick, D.J., Stirzaker, C., and Clark, S.J. (2012). Acetylation of H2A.Z is a key epigenetic modification associated with gene deregulation and epigenetic remodeling in cancer. Genome Research *22*, 307-321.

Voigt, P., Tee, W.-W., and Reinberg, D. (2013). A double take on bivalent promoters. Genes & Development *27*, 1318-1338.

Vongs, A., Kakutani, T., Martienssen, R., and Richards, E. (1993). Arabidopsis thaliana DNA methylation mutants. Science *260*, 1926-1928.

Wallace, J., and Orr-Weaver, T. (2005). Replication of heterochromatin: insights into mechanisms of epigenetic inheritance. Chromosoma *114*, 389-402.

Wang, A.Y., Aristizabal, M.J., Ryan, C., Krogan, N.J., and Kobor, M.S. (2011). Key Functional Regions in the Histone Variant H2A.Z C-Terminal Docking Domain. Molecular and Cellular Biology *31*, 3871-3884.

Watanabe, S., and Peterson, C.L. (2010). The INO80 Family of Chromatin-Remodeling Enzymes: Regulators of Histone Variant Dynamics. Cold Spring Harbor Symposia on Quantitative Biology *75*, 35-42.

Watanabe, S., Radman-Livaja, M., Rando, O.J., and Peterson, C.L. (2013). A Histone Acetylation Switch Regulates H2A.Z Deposition by the SWR-C Remodeling Enzyme. Science *340*, 195-199.

West, M.H.P., and Bonner, W.M. (1980). Histone 2A, a heteromorphous family of eight protein species. Biochemistry *19*, 3238-3245.

Williams, J.S., Hayashi, T., Yanagida, M., and Russell, P. (2009). Fission Yeast Scm3 Mediates Stable Assembly of Cnp1/CENP-A into Centromeric Chromatin. Molecular cell *33*, 287-298.

Wollmann, H., Holec, S., Alden, K., Clarke, N.D., Jacques, P.-â., and Berger, F.d.r. (2012). Dynamic Deposition of Histone Variant H3.3 Accompanies Developmental Remodeling of the <italic>Arabidopsis</italic> Transcriptome. PLoS Genet *8*, e1002658.

Wong, L.H., Ren, H., Williams, E., McGhie, J., Ahn, S., Sim, M., Tam, A., Earle, E., Anderson, M.A., Mann, J.*, et al.* (2009). Histone H3.3 incorporation provides a unique and functionally essential telomeric chromatin in embryonic stem cells. Genome Research *19*, 404-414.

Wratting, D., Thistlethwaite, A., Harris, M., Zeef, L.A.H., and Millar, C.B. (2012). A Conserved Function for the H2A.Z C Terminus. Journal of Biological Chemistry *287*, 19148-19157.

Wu, W.-H., Alami, S., Luk, E., Wu, C.-H., Sen, S., Mizuguchi, G., Wei, D., and Wu, C. (2005). Swc2 is a widely conserved H2AZ-binding module essential for ATP-dependent histone exchange. Nat Struct Mol Biol *12*, 1064-1071.

Xiao, A., Li, H., Shechter, D., Ahn, S.H., Fabrizio, L.A., Erdjument-Bromage, H., Ishibe-Murakami, S., Wang, B., Tempst, P., Hofmann, K.*, et al.* (2009). WSTF regulates the H2A.X DNA damage response via a novel tyrosine kinase activity. Nature *457*, 57-62.

Xiao, S., Xie, D., Cao, X., Yu, P., Xing, X., Chen, C.-C., Musselman, M., Xie, M., West, Franklin D., Lewin, Harris A.*, et al.* (2012). Comparative Epigenomic Annotation of Regulatory DNA. Cell *149*, 1381-1392.

Yelina, N.E., Choi, K., Chelysheva, L., Macaulay, M., de Snoo, B., Wijnker, E., Miller, N., Drouaud, J., Grelon, M., Copenhaver, G.P.*, et al.* (2012). Epigenetic Remodeling of Meiotic Crossover Frequency in <italic>Arabidopsis thaliana</italic> DNA Methyltransferase Mutants. PLoS Genet *8*, e1002844.

Yen, K., Vinayachandran, V., Batta, K., Koerber, R.T., and Pugh, B.F. (2012). Genome-wide Nucleosome Specificity and Directionality of Chromatin Remodelers. Cell *149*, 1461-1473.

Yen, K., Vinayachandran, V., and Pugh, B.F. (2013). SWR-C and INO80 Chromatin Remodelers Recognize Nucleosome-free Regions Near +1 Nucleosomes. Cell *154*, 1246-1256.

Zemach, A., Kim, M.Y., Hsieh, P.-H., Coleman-Derr, D., Eshed-Williams, L., Thao, K., Harmer, S.L., and Zilberman, D. (2013). The Arabidopsis Nucleosome Remodeler DDM1 Allows DNA Methyltransferases to Access H1-Containing Heterochromatin. Cell *153*, 193-205.

Zhou, J., Fan, J.Y., Rangasamy, D., and Tremethick, D.J. (2007). The nucleosome surface regulates chromatin compaction and couples it with transcriptional repression. Nat Struct Mol Biol *14*, 1070-1076.

Zilberman, D., Coleman-Derr, D., Ballinger, T., and Henikoff, S. (2008). Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. Nature *456*, 125-129.